Technische Hogeschool Eindhoven

# *Onderafdeling der Wiskunde*

# *Lineaire modellen*

naar het college van prof. dr. R. Doornbos

samengesteld door:
drs. A.J. Bosch en ir. W.L.M.M. Senden

Dictaatnr. 2.221 Prijs *f* 4,50

TECHNISCHE HOGESCHOOL EINDHOVEN
Afdeling Algemene Wetenschappen
Onderafdeling der Wiskunde

# LINEAIRE MODELLEN

naar het college van

**Prof. Dr. R. Doornbos**

samengesteld door

**Drs. A.J. Bosch**

en

**Ir. W.L.M.M. Senden**

Najaarssemester 1976

TECHNISCHE HOGESCHOOL EINDHOVEN

Onderafdeling der Wiskunde

# LINEAIRE MODELLEN

naar het college van prof. Dr. R. Doornbos

samengesteld door:

Drs. A.J. Bosch en Ir. W.L.M.M. Senden

Najaarssemester 1976

# Inhoudsbeschrijving

## LINEAIRE MODELLEN

## R. Doornbos, A.J. Bosch en W.L.M.M. Senden

### Najaarssemester 1976



JdG, 25 Juli 2005

Inhoud



Literatuur:

[1] Rao, C.R., Linear statistical inference and its applications, Wiley 1965.

[2] Searle, S.R., Linear models, Wiley 1971.

[3] Seber, G.A.F., The linear hypothesis, A general theory, Griffin 1966.

## § 0. Inleiding

De stof van § 4 t/m 11 en appendix D is ontleend aan Searle [2], zij het in een andere volgorde en met enigszins andere notaties.
Om deze stof goed te kunnen behandelen is een zekere basiskennis nodig § 1 t/m 4; tevens is appendix D van groot belang.

Voor het bestuderen van dit dictaat kan men de volgende punten overslaan:
uit § 1 : 1.5; 1.6; 1.23; 1.59; 1.69; 1.70; 1.78;
1.79; 1.98; 1.99; 1.100;
uit § 2 : 2.17; 2.45; 2.46;
Uit appendix B : B3; B5; B7; geheel Bc.

## § 1. Enkele punten uit de matrixrekening

Voor de behandeling van de multivariate analyse, lineaire modellen e.d. is de theorie der lineaire vectorruimten en de matrixrekening van fundamenteel belang. Met behulp van matrices worden de notaties veel overzichtelijker en voor vele problemen geeft een vectormodel een beter inzicht.
Vooral de orthogonale projectie blijkt een belangrijke rol te spelen.

### a. Definities, notaties

1.1. $A_{n\times p} = \{a_{ij}\}$ is een *matrix* van de *orde* $n\times p$ met n rijen, p kolommen en elementen $a_{ij}$; $(A)_{ij} = a_{ij}$.
$A_{n\times n}$ is een *vierkante* matrix; $A_{n\times n}$ is een *diagonaal*matrix als $a_{ij} = 0$ voor $i \neq j$.
$A_{n\times n} = \mathrm{diag}(a_{11},\ldots,a_{nn})$ d.w.z. A is een diagonaalmatrix met diagonaalelementen $a_{11},\ldots,a_{nn}$. Zo ook $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_n)$.
$I_{n\times n} =: I_n = \{\delta_{ij}\}$ is de *eenheids*matrix, $\delta_{ij}$ het Kronecker-symbool.
$A = \alpha I$ is een *scalaire* matrix.
$A'$ is de *getransponeerde* matrix van A d.w.z. $(A')_{ij} = a_{ji}$ voor alle i,j.
In het bijzonder is $a'$ een *rijvector* en a de bijbehorende *kolomvector*.
$A_{n\times n}$ heet *symmetrisch* als $A' = A$ (ook wel $A^T$ i.p.v. $A'$ geschreven).
$A_{n\times n}$ heet *antisymmetrisch* als $A' = -A$.
$u = (1,\ldots,1)'$ ; $U = \{1\}$ (in de literatuur vaak J, matrix van enen).

$a_{i*}$ is de i-de rij van A als kolom geschreven,

$a'_{i*}$ is de i-de rij van A als rij geschreven,

$a_{*j}$ is de j-de kolom van A als kolom geschreven,

$a'_{*j}$ is de j-de kolom van A als rij geschreven.

Zo is $(AB)_{ij} = a'_{i*}b_{*j}$.
$A_{p\times q} = (a_{*1} \ \ldots \ a_{*q})$ ; $I_p = (\delta_{*1} \ \ldots \ \delta_{*p})$.

$\tilde{I}_p := (\delta_{*i_1} \ \ldots \ \delta_{*i_p})$ d.w.z. de kolommen van $I_p$ in andere volgorde.

Deze matrix heet een *permutatie*matrix.

Opmerking: Daar we gericht zijn op de statistiek, worden alle gegeven grootheden reëel ondersteld. Ook zullen we geen onderscheid maken in notatie tussen een lineaire afbeelding $\mathcal{A}$ en de bijbehorende matrix A, daar we steeds met één basis werken die bovendien orthonormaal ondersteld is.

1.2. $|A|$ = det A is de *determinant* van A; abs$|A|$ is de absolute waarde van det A.
$|A|_{ij}$ is de *cofactor* van $a_{ij}$, dat is $(-1)^{i+j}$ × de onderdeterminant van $a_{ij}$.
$A^{-1}$ is de *inverse* van $A_{n\times n}$ zo deze bestaat.
$(A)^{ij} := (A^{-1})_{ij}$, ook wel als $a^{ij}$ genoteerd.

1.3. sp $A_{n\times n} := \sum_1^n a_{ii}$ heet het *spoor* van A (Engels: trace).

1.4. r(A) is de *rang* van A, dat is het maximum aantal onafhankelijke rijen van A.
$A_{n\times p}$ heet van *volle rang* als r(A) = min(n,p).
$A_{n\times p}$ heet van *volle rijrang* als r(A) = n (dus $n \le p$).
$A_{n\times p}$ heet van *volle kolomrang* als r(A) = p (dus $p \le n$).
$A \in \mathcal{M}^r_{p,q}$ d.w.z. A is een matrix van de orde p×q met rang r.

1.5. Het *direct-* of *Kronecker product* van twee willekeurige matrices A en B is
$A \otimes B := \{a_{ij}B\}$, d.w.z. elk element van A wordt met B vermenigvuldigd.

1.6. $\vec{A} := (a'_{*1} \ \dots \ a'_{*q})'$, d.w.z. de kolommen van A achtereenvolgens onder elkaar genoteerd (ook wel *vec* A genoemd). Als A van de orde p×q is, is $\vec{A}$ dus een vector met pq componenten.

1.7. We definiëren voor elke matrix A een "*pseudokwadraat*", nl. $A^{\underline{2}} := A'A$.
Analoog $a^{\underline{2}} := a'a$ ; $A^{\underline{-2}} := (A^{\underline{2}})^{-1}$.
Als het geen verwarring schept, schrijven we vaak $A^2$ i.p.v. $A^{\underline{2}}$ (bijv. als A niet vierkant is bestaat het gewone kwadraat $A^2$ niet).
In de multivariate-statistiek komt het pseudodwadraat echter veel vaker voor, het gewone vrijwel niet. Zoals we zullen zien biedt het nog meer voordelen.

1.8. Het *inwendig product* van de vectoren a en b wordt genoteerd als (a,b). Daar we steeds (vanaf nu stilzwijgend) een orthonormale basis veronderstellen, is $(a,b) = a'b = b'a = \Sigma a_i b_i$. Opmerking: vectoren onderstrepen we niet. Analoog voor matrices A en B van dezelfde orde: $(A,B) := sp(A'B)$. Eenvoudig is na te gaan dat dit inderdaad een inproduct definiëert.

1.9. Via het inproduct worden gedefiniëerd:

De *norm* van een vector a is $\|a\| := \sqrt{(a,a)} = \sqrt{a^2}$ ; analoog

de *norm* van een matrix A is $\|A\| := \sqrt{(A,A)} = \sqrt{spA^2}$ .

De *hoek* $\psi$ tussen 2 vectoren a en b (beide $\neq 0$) : $\cos\psi := \dfrac{(a,b)}{\|a\|\,\|b\|}$ ;

a en b heten *orthogonaal*, notatie $a \perp b$, als $(a,b) = 0$.
Een reële vectorruimte met inproduct wordt *Euclidische ruimte* $E_n$ genoemd.

$\lim\limits_{n\to\infty} A_n = A$, d.w.z. $\|A_n - A\| < \varepsilon$ voor $n > N$.

Zij $a \in V$, $B \subset V$; $d[a,B] := \inf\limits_b \{\|a-b\|;\ b \in B\}$ heet de *afstand* van a tot B.

Opmerking: dat deze norm gekozen is hangt samen met een metriek, waarbij de methode der kleinste kwadraten wordt toegepast. Er zijn andere normen mogelijk.

1.10. $A_{n\times n}$ heet *regulier* als $|A| \neq 0$, anders *singulier*.

1.11. $A_{n\times n}$ heet *idempotent* als $AA = A$ (dus ook als $A^2 = A$).

1.12. Een reguliere matrix A heet *orthogonaal* als $A^2 = I$ oftewel $A' = A^{-1}$.

1.13. $x'Ax$ is een *kwadratische vorm* in x. Hierbij wordt A symmetrisch ondersteld (zie 1.56).

1.14. Een symmetrische matrix A heet *positief definiet* (afgekort p.d. of $A > 0$) als $\forall x \neq 0$ $(Ax,x) = x'Ax > 0$.
Is $x'Ax \geq 0$ voor alle x en $\exists x \neq 0$ met $x'Ax = 0$, dan heet A *semipositief definiet*. Is $x'Ax \geq 0$ voor alle x, dan is A *niet negatief definiet* $(A \geq 0)$.
$A > B$ d.w.z. $A - B > 0$.

1.15. Zij V een reële vectorruimte. Als $\mathcal{A}x = \lambda x$ voor $x \neq 0 \in V$ en $\lambda$ reëel, dan heet x *eigenvector* en $\lambda$ de bijbehorende *eigenwaarde* van de afbeelding $\mathcal{A}$; de eigenwaarden van $\mathcal{A}$ zijn de <u>reële</u> wortels van de *karakteristieke vergelijking* $|A-\lambda I| = 0$. De eigenwaarden van de <u>matrix</u> A zijn <u>alle</u> wortels, die dus complex kunnen zijn.

1.16. Een matrix A heet *defect* als er gēēn reguliere matrix B bestaat zodat $B^{-1}AB = \Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_p)$; B en $\Lambda$ mogen complexe elementen bevatten. Zijn B en $\Lambda$ reëel, dan heet A *reëel diagonaliseerbaar*.

1.17. Het stelsel lineaire vergelijkingen $Ax = b$ heet *consistent* als er minstens een oplossing is.

1.18. $N(A) := \{x \in V | Ax = 0\}$ heet de *nulruimte* of *kern* van A.

1.19. $\langle A\rangle = \langle a_{*1} \ldots a_{*q}\rangle$ is de *kolommenruimte* van $A_{p\times q}$, dat is de lineaire deelruimte in $R^p$ opgespannen door de kolommen van A.
Analoog is $\langle A'\rangle$ de *rijenruimte* van A, dat is de lineaire deelruimte in $R^q$ opgespannen door de rijen van A.

1.20. $E_\lambda(\mathcal{A}) := \{x \in V | \mathcal{A}x = \lambda x\}$ is de *eigenruimte* van $\mathcal{A}$ behorende bij de eigenwaarde $\lambda$; dim $E_\lambda$ heet de *meetkundige multipliciteit* van $\lambda$.

1.21. Zij $D \subset V$ een lineaire deelruimte van V en $a \in V$.
$S = a+D$ heet een *lineaire variëteit* en dim S := dim D.

1.22. Zij D een lineaire deelruimte van de Euclidische ruimte $E_n$, dan is $D^\perp := \{x \in E_n | x \perp D\}$ het *orthogonale complement* (*orthoplement*) van D in $E_n$.

1.23. Het *spectrum* $\sigma(A)$ van A is de verzameling van alle eigenwaarden van A.

b. Eigenschappen

1.30. Een belangrijk hulpmiddel is het *partitioneren* van matrices:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} E & F \\ G & H \end{pmatrix} = \begin{pmatrix} AE + BG & AF + BH \\ CE + DG & CF + DH \end{pmatrix} .$$

De afmetingen zijn zodanig dat alle produkten bestaan.
In het bijzonder is: $AB = \sum_j a_{*j}\, b'_{j*}$; $Ax = b$ geeft $\sum_j a_{*j}\, x_j = b$.

$A'A = A^{\underline{2}} = \sum_1^n a_{i*} a'_{i*}$. $\quad (A \in \mathcal{M}_{p,n})$ .

Analoog: $AA' = (A')^{\underline{2}} = \sum_1^p a_{*j}\, a'_{*j}$.

1.31. In het algemeen is $AB \neq BA$ (dus ook $A'A \neq AA'$). Is echter $B = \beta I$, dan is $AB = BA$; tevens is $B' = B$.

1.32. sp is een *lineaire operator*, nl. $sp(\alpha A + \beta B) = \alpha\, sp\, A + \beta\, sp\, B$.
Tevens geldt $sp(AB) = sp(BA)$. Bewijs:

$$sp(AB) = \sum_i (AB)_{ii} = \sum_{i,k} a_{ik}\, b_{ki} = \sum_{k,i} b_{ki}\, a_{ik} = \sum_k (BA)_{kk} = sp(BA) .$$

In het bijzonder geldt: $sp(ABC) = sp(CAB) = sp(BCA)$

$$sp(ab') = sp(b'a) = b'a;\ sp\, A = sp\, A';$$

$$\|A\|^2 = sp\, A^{\underline{2}} = sp \sum_i a_{i*}\, a'_{i*} = sp \sum_i a'_{i*}\, a_{i*} = \sum_i a^2_{i*} = \sum_{ij} a^2_{ij} .$$

Oftewel:
De norm van een matrix is de wortel uit de som van de kwadraten der elementen. Bovendien is $sp\, A = \Sigma\lambda_i$ (*som der eigenwaarden* van A).

1.33. $(AB)^{-1} = B^{-1}A^{-1}$ (A,B n×n matrices en regulier)

$$|A| = \sum_i a_{ij}|A|_{ij} = \sum_j a_{ij}|A|_{ij} \ ; \ |A^{-1}| = 1/|A| \ ; \ |AB| = |A|\,|B| \ ;$$

$$a^{ij} = (A^{-1})_{ij} = |A|_{ji}/|A| .$$

Bovendien is $|A| = \Pi\lambda_i$ (*product der eigenwaarden* van A).
Daar $|A-\lambda I| = \Pi(\lambda_i-\lambda)$ is, neem $\lambda = 1$: $|A-I| = \Pi(\lambda_i-1) = |\Lambda-I|$.

1.34. $\begin{vmatrix} A & B \\ O & C \end{vmatrix} = |A|\,|C|$ (A,C vierkante matrices) nl.

$$\begin{vmatrix} A & B \\ O & C \end{vmatrix} = \begin{vmatrix} I & O \\ O & C \end{vmatrix}\begin{vmatrix} A & B \\ O & I \end{vmatrix} = |C|\,|A| \ .$$

1.35. Is $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ ($A_{11}$, $A_{22}$ vierkant, $A_{22}$ regulier), dan geldt:

$$|A| = \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix}\begin{vmatrix} I_{k\times k} & 0 \\ -A_{22}^{-1}A_{21} & I_{r\times r} \end{vmatrix} = \begin{vmatrix} A_{11\cdot 2} & A_{12} \\ 0 & A_{22} \end{vmatrix} = |A_{11\cdot 2}|\,|A_{22}|$$

waarin $A_{11\cdot 2} := A_{11} - A_{12}A_{22}^{-1}A_{21}$.
Analoog geldt $|A| = |A_{22\cdot 1}|\,|A_{11}|$ (nu alleen $A_{11}$ regulier ondersteld).
Deze enigszins vreemde eigenschap wordt gebruikt bij het afleiden van toetsingsgrootheden via het likelihood-ratio-criterium.

1.36. Zij $A \in \mathcal{M}^r_{p,q}$ en gepartitioneerd als in 1.35 met $r(A_{11}) = r$.
Dan is $A_{22.1} = 0$.
Bewijs:
Daar $r(A) = r(A_{11})$ is $(A_{21}A_{22}) = K(A_{11}A_{12})$; oftewel $A_{21} = KA_{11}$ dus
$K = A_{21}A_{11}^{-1}$ en $A_{22} = KA_{12} = A_{21}A_{11}^{-1}A_{12}$ (dus $A_{22.1} = 0$).

1.37, Is A orthogonaal, dan is $(Ax,Ay) = (x,y)$ nl. $(Ax,Ay) = x'A'Ay = x'y = (x,y)$.
In het bijzonder is $\|Ax\| = \|x\|$.
$A^{-1}$ is eveneens orthogonaal daar $(A^{-1})' = (A')^{-1} = (A^{-1})^{-1}$ (zie 1.51).

1.38. $\tilde{I}_p^2 = I_p$, dus orthogonaal, nl. $(\tilde{I}^2)_{ij} = \delta'_{*i_i}\,\delta_{*i_j} = \delta_{ij}$.
Zo is $A_{p\times q}\tilde{I}_q$ een matrix met dezelfde kolommen als van A, alleen gepermuteerd.
En $\tilde{I}_p A_{p\times q}$ " " " " rijen " " ", " "
Wil men de <u>kolommen</u> van A als volgt hebben: $\tilde{A} = (a_{*3}\ a_{*4}\ a_{*1}\ a_{*2})$ dan moet men navermenigvuldigen met $\tilde{I}_4 = (\delta_{*3}\ \delta_{*4}\ \delta_{*1}\ \delta_{*2})$.

1.39. Zij $A \in L(V,W)$, d.w.z. een lineaire afbeelding van V in W, dan geldt:

$$\dim N(A) + r(A) = \dim V.$$

Hieruit volgt:

Zij $A \in \mathcal{M}^r_{p,q}$ en $Ax = b$ consistent, dan is dim. oplossingsvariëteit S

$\dim S = q-r$ . (Zie Wsk 20.)

1.40. Zij D een lineaire deelruimte $\subset E_n$, dan geldt:

$\dim D + \dim D^\perp = n$ (Zie Wsk 30.)

1.41. $N(A') = \langle A\rangle^\perp$ . (Zie Wsk 30.)

1.42. Zij $A \in \mathcal{M}^{q-m}_{p,q} \Rightarrow \exists D \in \mathcal{M}^m_{q,m}$ met $AD = 0$ nl.

$\dim N(A) = m$ dus $N(A) = \langle d_{*1} \dots d_{*m}\rangle$ en $A(d_{*1} \dots d_{*m}) =: AD = 0$.
Tevens geldt:

Stel $B \in \mathcal{M}^m_{m,q}$ ; BD regulier $\Leftrightarrow \langle B'\rangle \cap \langle A'\rangle = \{\vec{o}\}$,

oftewel rijen van B onafhankelijk rijen van A.

Bewijs:

$\Rightarrow$ Stel $\exists b_{i*} \in \langle A'\rangle$, d.w.z. $\exists t$ met $A't = b_{i*}$ dus $D'b_{i*} = D'A't = \vec{o}$.
Nu is BD, dus D'B' regulier en $D'b_{i*} = (D'B')_{*i} = o$ strijdig.

$\Leftarrow$ Nu is $\binom{A}{B} \in \mathcal{M}^q_{p+m,q}$ . Volgens 1.80 is er een linkerinverse (VW) zodat $VA + WB = I_q$. Dus $VAD + WBD = D$ oftewel $WBD = D$.
Analoog is er een linkerinverse E zodat $ED = EWBD = I_m$.
Nu is $m = r(EWBD) \le r(BD) \le m$ dus $r(BD) = m$ oftewel BD regulier.

1.43. Zij $y_{q\times 1} \neq \vec{o}$, $z_{p\times 1}$ gegeven vectoren. Dan is er een $X_{p\times q}$ met $Xy = z$, nl. Stel $y_k \neq o$. Neem $x_{*j} = o$ voor $j \neq k$ en $x_{*k} = z/y_k$, dan is

$$Xy = \sum_j x_{*j}\, y_j = z.$$

## De rang van een matrix

Nu volgen enkele eigenschappen betreffende de rang van een produkt van matrices. Dit is belangrijk o.a. om te kijken of een produkt een inverse heeft en om de dimensie van de beeldruimte te bepalen.

1.44. a) $r(A') = r(A)$ (zie Wiskunde 10).

b) $r(AB) \le \min[r(A), r(B)]$, nl. $r(AB) = r(Ab_{*1}, \ldots, Ab_{*s}) =$
$= r(\Sigma a_{*i} b_{i1}, \ldots, \Sigma a_{*i} b_{is}) \le r(A)$.
Analoog: $r(AB) = r(B'A') \le r(B') = r(B)$. Dus $r(ab') = 1$ als $a,b \neq 0$.

c) $A \in \mathfrak{M}^{q}_{p,q}$: $r(AB) = r(B)$, nl.: $r(B) \ge r(AB) = r(B'A') \ge r[B'A'A(A'A)^{-1}] =$
$= r(B') = r(B)$; $A'A$ is regulier volgens 1.45 nl. $r(A'A) = r(A) = q$.

$B \in \mathfrak{M}^{q}_{q,s}$: $r(AB) = r(A)$, nl.: $r(AB) = r(B'A') = r(A') = r(A)$.
Bijzonder geval: de rang verandert niet door voor- of navermenigvuldiging met <u>reguliere</u> matrix.

1.45. $r(A^2) = r(A)$ (analoog $r(AA') = r(A)$; in het algemeen is $r(AB) \neq r(BA)$).

<u>Bewijs</u>:
$Ax = o$ geeft $A^2x = o$ dus $N(A) \subset N(A^2)$; $A^2x = o$ geeft $x'A'Ax = (Ax)^2 = o$
dus $Ax = o$ oftewel $N(A^2) \subset N(A)$; $\dim N(A^2) = \dim N(A)$, dus $r(A^2) = r(A)$.

1.46. Zij $r(A) = r$, dan bevat $A$ een <u>reguliere</u> $r \times r$ ondermatrix (zie Wsk 10).

1.47. Als $ABA = A \Rightarrow r(AB) = r(BA) = r(A)$.

<u>Bewijs</u>:
$r(A) \le r(AB)$ èn $r(AB) \le r(A)$ dus $r(AB) = r(A)$.

1.48. $\langle AB\rangle \subset \langle A\rangle$, nl. $\langle AB\rangle = \langle Ab_{*1} \ldots Ab_{*q}\rangle \subset \langle A\rangle$.
Is $B$ regulier of van volle <u>rij</u>rang (zie 1.44) dan is $\langle AB\rangle = \langle A\rangle$.
Verder is $\langle X'X\rangle = \langle X'\rangle$ (met 1.45).

1.49. $Ax = b$ consistent $\Leftrightarrow b \in \langle A\rangle$. Ga zelf na (Wsk 10).

<u>Gevolg</u>: Is $A \in \mathcal{M}^{p}_{pq}$ dan is $Ax = b$ consistent voor <u>alle</u> $b_{p\times 1}$.

1.50. $X'Xb = X'y$ is consistent daar $X'y \in \langle X'\rangle = \langle X'X\rangle$.

Symmetrische en positief definiete matrices

Nu volgen enkele eigenschappen betreffende symmetrische en positief definiete matrices. Deze spelen een belangrijke rol o.a. omdat de variantie-covariantie matrix (semi)positief definiet is.

1.51. $(A+B)' = A' + B'$; $(AB)' = B'A'$.

Bewijs:

$(AB)'_{ij} = (AB)_{ji} = \sum_k a_{jk}b_{ki} = \sum_k b'_{ik}\, a'_{kj} = (B'A')_{ij}$.

$(A^{-1})' = (A')^{-1}$ nl. $AA^{-1} = I$ geeft $(A^{-1})'A' = I' = I$.

Gevolg: A regulier, symmetrisch $\Longleftrightarrow$ $A^{-1}$ regulier, symmetrisch.

Is $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ dan is $A' = \begin{pmatrix} A'_{11} & A'_{21} \\ A'_{12} & A'_{22} \end{pmatrix}$.

1.52. $\forall x,y\ (Ax,y) = (x,A'y)$, nl. $(Ax,y) = x'A'y = (x,A'y)$.
Is A symmetrisch, dan geldt $(Ax,y) = (x,Ay)$.
Opmerking: Dit is eigenlijk de definitie van een symmetrische afbeelding. Op een orthonormale basis behoort bij een symmetrische afbeelding een symmetrische matrix en omgekeerd. Ditzelfde geldt voor een orthogonale afbeelding en een orthogonale matrix.

1.53. Is A symmetrisch en $\lambda_1 \neq \lambda_2$, dan zijn de bijbehorende eigenvectoren $x_1$ en $x_2$ orthogonaal, nl. $\lambda_1(x_1,x_2) = (Ax_1,x_2) = (x_1,Ax_2) = \lambda_2(x_1,x_2)$ dus $(x_1,x_2) = o$.

1.54. Uit 1.16 volgt voor een niet defecte matrix A : er is een reguliere matrix B zodat $AB = B\Lambda$ oftewel $A = B\Lambda B^{-1}$; B bestaat uit eigenvectoren en $\Lambda = \mathrm{diag}(\lambda_1,\ldots,\lambda_p)$ uit de bijbehorende eigenwaarden van A.
Nu geldt: een symmetrische matrix A is reëel diagonaliseerbaar (zie Wsk 30). Met 1.53 kan men er voor zorgen dat B orthogonaal is, zodat

$$A = H\Lambda H'$$

(H nu een orthogonale matrix van eigenvectoren en $\Lambda$ de diagonaalmatrix van eigenwaarden; in het vervolg zullen H en $\Lambda$ steeds deze betekenis hebben).

1.55. Zij $A \in \mathcal{M}^{p-m}_{p,p}$, symmetrisch en $0 < m < p$. Dan is er een $B \in \mathcal{M}^{m}_{m,p}$ zodat $\begin{pmatrix} A & B' \\ B & 0 \end{pmatrix}$ regulier is.

<u>Bewijs:</u>

$r(A) = p-m$, dus $\langle A^T \rangle = \langle A \rangle$ aan te vullen met m onafhankelijke rijen zodat $r\binom{A}{B} = p$. Oftewel voor alle i geldt $b_{i*} \notin \langle A \rangle$ dus ook

$$\begin{pmatrix} b_{i*} \\ o \\ \vdots \\ o \end{pmatrix} \notin \left\langle \begin{matrix} A \\ B \end{matrix} \right\rangle \quad \text{oftewel } r\begin{pmatrix} A & B' \\ B & 0 \end{pmatrix} = p+m \text{ dus regulier.}$$

1.56. $f(x) = x'Bx = \frac{1}{2}(x'Bx+x'B'x) = x'(\frac{B+B'}{2})x =: x'Ax = f(x)$ nu echter met A <u>symmetrisch</u>. Vandaar dat in een kwadratische vorm de matrix zonder beperking symmetrisch ondersteld kan worden, deze is bovendien uniek (met 1.57).

1.57. $\forall x,y : x'Ay = o \iff A = 0$, nl. $\delta'_{*i}\, A\delta_{*j} = a_{ij} = o\ \forall ij$.
$\forall x \quad : x'Ax = o \iff A = -A'$ (A nu niet symmetrisch ondersteld!).

<u>Bewijs:</u> $\Leftarrow x'Ax = \frac{1}{2}x'(A+A')x = o$

$\Rightarrow \delta'_{*i}\, A\delta_{*i} = a_{ii} = o$. Nu $(o,\ldots,1,\ldots,1,\ldots,o)\, A \begin{pmatrix} o \\ \vdots \\ 1 \\ \vdots \\ 1 \\ \vdots \\ o \end{pmatrix} = a_{ij} + a_{ji} = o$

geeft $a_{ij} = -a_{ji}$.

Opmerking: is A symmetrisch dan volgt dat $A = 0$.

1.58. $A^{\underline{2}} = 0 \iff A = 0$, nl. $r(A^{\underline{2}}) = r(A) = o$ dus $A = 0$,
in het bijzonder $\|A\| = o \iff A = 0$.

1.59. $A^{\underline{2}} = A^2 \iff A = A'$, nl.
$\|A'-A\|^2 = sp(A-A')(A'-A) = sp[AA'-AA-A'A'+A'A] = 2sp(A'A) - 2sp(AA) = o$
daar $sp(A^{\underline{2}}) = sp(A^2)$, dus met (1.58) $A = A'$.

1.60. $AX'X = BX'X \iff AX' = BX'$, nl.
$(A-B)X'X = 0$; $\langle (A-B)' \rangle \subset \langle X'X \rangle^{\perp} = \langle X' \rangle^{\perp}$ dus $(A-B)X' = 0$. Ander bewijs:
$0 = (A-B)X'X(A-B)' = [X(A-B)']^{\underline{2}}$ dus $X(A-B)' = (A-B)X' = 0$ (1.58).

Opmerking: $A \in \mathcal{M}_{pq}^{q}$ (dus van volle kolomrang) dan $AB = AC \Longleftrightarrow B = C$, nl.
$A'AB = A'AC$ en daar $A'A$ nu regulier is (met 1.45) volgt $B = C$.
Analoog $A \in \mathcal{M}_{pq}^{p}$ (dus van volle rijrang) dan $BA = CA \Longleftrightarrow B = C$.

1.61. $A,B \geq o \Rightarrow A+B \geq o$, nl. $((A+B)x,x) = (Ax,x) + (Bx,x) \geq o$.
Is $N(A) \cap N(B) = \{o\}$ dan is $A+B > o$; in het bijzonder is $A+I > o$.

1.62. $A'B = 0 \Rightarrow (A+B)^{\underline{2}} = A^{\underline{2}} + B^{\underline{2}}$ (niet omgekeerd), nl.
$(A+B)'(A+B) = (A'+B')(A+B) = A'A + B'B + A'B + B'A = A^{\underline{2}} + B^{\underline{2}}$.

1.63. Zij $A$ symmetrisch, dan geldt $A > o \Leftrightarrow \forall i\ \lambda_i > o$.

Bewijs: Volgt direct uit 1.66.

1.64. $A \geq 0 \Longleftrightarrow A = D^{\underline{2}}$.

Bewijs $\Rightarrow$ : neem $D = \Lambda^{\frac{1}{2}}H'$, dan is $D^{\underline{2}} = H\Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}H' = H\Lambda H' = A$.

Bewijs $\Leftarrow$ : $(Ax,x) = (D^{\underline{2}}x,x) = (Dx,Dx) \geq o$.

Opmerking : Is $A > 0$ dan is $D$ van volle kolomrang en omgekeerd.
In de gebruikelijke notatie $D^2 = DD$ is dit niet waar. Neem bijv. $D = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$
$DD = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, sp $DD < 0$ oftewel met (1.63) is $DD$ niet $> 0$.
Ook geldt $A = BB$, nl. met $B := H\Lambda^{\frac{1}{2}}H'$; $B =: A^{\frac{1}{2}}$.
Is $A > 0$ dan is er zelfs een bovendriehoeksmatrix $D = \begin{pmatrix} \ddots & \vdots \\ 0 & \ddots \end{pmatrix}$ met $A = D^{\underline{2}}$.

Deze is te vinden door een *Gram-Schmidt orthogonalisatieproces*.
$D$ is uniek als men eist $d_{ii} \geq o$ voor alle $i$ (*Choleskiprocedure*).

1.65. $A > 0 \Longleftrightarrow A^{-1} > 0$.

Bewijs: volgt direct uit 1.66.

1.66. $B \in \mathfrak{M}^q_{q,p}$: $A_{p\times p} > 0 \Rightarrow BAB' > 0$.

Bewijs: $BAB'$ is symmetrisch en $(BAB'x,x) = (AB'x,B'x) = (Ay,y) > 0$ daar $y \neq 0$ als $x \neq 0$.

Opmerking:

a) Is slechts gegeven dat A regulier is, dan kan BAB' wel singulier zijn.

b) Is B regulier dan geldt: $A > 0 \Leftrightarrow BAB' > 0$.

c) Daar $A = H\Lambda H'$ geldt dus ook: $A > 0 \Leftrightarrow \Lambda > 0$ (dus alle $\lambda_i > 0$).

d) Daar $|A| = \Pi\lambda_i$, is A regulier.

e) $A > 0 \Leftrightarrow \Lambda > 0 \Leftrightarrow \Lambda^{-1} > 0 \Leftrightarrow A^{-1} > 0$ ($A^{-1}$ is ook symmetrisch en bestaat).

1.67. $A > 0 \Longleftrightarrow$ elke hoofdondermatrix $A_{ij} > 0$. Stel $A_{p\times p} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$.

Zij $A_{11}$ van de orde $p_1 \times p_1$, $A_{21}$ van de orde $p_2 \times p_1$, dus $p_1 + p_2 = p$.
Neem $x' = (x_1', o')$ waarin $x_1$ een vector met $p_1$, o een vector met $p_2$ componenten is. Voor alle $x_1 \neq o$ is $x_1'A_{11}x_1 = x'Ax > 0$ dus met A ook $A_{11} > 0$.
Een willekeurige hoofdondermatrix B is altijd linksboven te plaatsen, nl $\tilde{A} = \tilde{I}'A\tilde{I}$. Daar I regulier is, is (met 1.66) $\tilde{A}$ ook $> 0$.

1.68. A,B symmetrisch en $A > 0$ dan is AB reëel diagonaliseerbaar.

Bewijs:
$A = D^2$ (1.64) dus $AB = D'DB = D'DBD'D^{-T}$ (A, dus D regulier).
Nu is $DBD'$ symmetrisch dus $= H\Lambda H^{-1}$ oftewel $AB = D'H\Lambda H^{-1}D^{-T} =: C\Lambda C^{-1}$ met $C := D'H$.
Dit is niet juist als A s.p.d. is!

1.69. Is $A \neq 0$ reëel diagonaliseerbaar dan geldt $sp(AA) > 0$.

Bewijs:
$A = B\Lambda B^{-1}$; $sp(AA) = sp(B\Lambda B^{-1}B\Lambda B^{-1}) = sp(\Lambda^2) > 0$.

Opmerking: is A defect dan is dit niet waar, voorbeeld:

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \; ; \; AA = 0 \text{ en } sp(AA) = 0.$$

1.70. Is A niet defect en $\lambda_i = \lambda$ voor alle i, dan is $A = \lambda I$, nl. $A = B\Lambda B^{-1} = \Lambda = \lambda I$.
Deze triviale uitspraak biedt de mogelijkheid defecte matrices te construeren, voorbeeld:

$$\begin{pmatrix} 3 & 0 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

1.71. Is A niet defect dan is $r(A) = \Sigma m_i$, waarin $m_i$ de algebraïsche multipliciteit is van $\lambda_i \neq o$; nl. $r(A) = r(B\Lambda B^{-1}) = r(\Lambda) = \Sigma m_i$.

1.72. Gegeven $A_{p\times q}$ met $p \leq q$. Dan zijn de eigenwaarden $\neq 0$ van $A'A$ dezelfde als die van $AA'$.

Bewijs:
Stel $A'Ax = \lambda x$ en $\lambda \neq o$, dus ook $(AA')Ax = \lambda Ax$ oftewel $\lambda$ eigenwaarde van $AA'$ (nl. $Ax \neq o$). Analoog als $\mu$ eigenwaarde is van $AA'$, dan ook van $A'A$.

1.73. Stel $A, B > 0$ en van dezelfde orde. Dan is er een reguliere D zodat $A = D^{\underline{2}}$ en $B = D'\Lambda D$ ($\Lambda$ diagonaal met alle $\lambda_i > o$).

Bewijs:
$A = E^{\underline{2}}$ (zie 1.64). Beschouw $E^{-T}BE^{-1}$ (volgens 1.66 ook $> 0$) dus $= H\Lambda H'$ ($\lambda_i > o$). Dus $B = E'H\Lambda H'E =: D'\Lambda D$ met $D = H'E$. Dit geeft $A = E'E = E'HH'E = D'D = D^{\underline{2}}$.
Opmerking: is B alleen symmetrisch, dan geldt de stelling ook, alleen vervalt $\lambda_i > 0$.

1.74. Stel $A > 0$, $B \geq 0$ en $B \neq 0$. Dan is $|A + B| > |A| + |B|$.

Bewijs:
Volgens 1.73 is

$$|A+B| = |D'D+D'\Lambda D| = |D'|\ |I+\Lambda|\ |D| = |A|\ |I+\Lambda| > |A|\ |\ (1+|\Lambda|) =$$

$$= |A| + |A|\ |\Lambda| = |A| + |B| \quad \text{nl.}$$

$$|I+\Lambda| = \Pi(1+\lambda_i) > 1 + \Pi\lambda_i = 1 + |\Lambda|.$$

Gevolg: Is $A > B \Rightarrow |A| > |B|$.

1.75. $sp(AA) \leq sp(A'A)$.

Bewijs:
$sp(A'\pm A)'(A'\pm A) \geq 0$ (zie 1.64 en 1.63)
$= sp[(A\pm A')(A'\pm A)] = sp[AA'+A'A\pm AA\pm A'A'] = 2sp[A'A\pm AA]$.
Dus $sp(A'A) \geq \pm sp(AA)$ of anders geschreven: $|sp\ AA| \leq \|A\|^2$.

1.76. $\|AB\|^2 \leq \|A\|^2\ \|B\|^2$.

Bewijs:
$$\|AB\|^2 = \sum_{ij} [(AB)_{ij}]^2 = \sum_{ij} (a'_{i*}b_{*j})^2 \leq \sum_{ij} a^2_{i*}b^2_{*j} \quad \text{(Cauchy-Schwarz)}$$
$$= \sum_i a^2_{i*} \sum_j b^2_{*j} = \|A\|^2\ \|B\|^2 .$$

Speciaal geval: $\|Ax\|^2 \le \|A\|^2 \|x\|^2 = \|A\|^2 x^2$ .

<u>Opmerking</u>: Cauchy-Schwarz ongelijkheid luidt: $(x,y)^2 \le (x,x)(y,y)$ oftewel $(x'y)^2 \le \|x\|^2 \|y\|^2 = x^2 y^2$; we zien dat $\|Ax\| \le M\|x\|$ voor alle x.

1.77. $\sum_0^\infty A^k = (I-A)^{-1}$ voor $\|A\| < 1$. (I - A is regulier, zie som 1.24).

<u>Bewijs</u>:

Definitie: $\lim_{n \to \infty} S_n = S$ dus $\|S_n - S\| < \varepsilon$ voor $n > N$.

Stel $S_n := \sum_0^n A^k$ dus $I-A^n = S_n(I-A)$ (som eindige meetkundige reeks met reden A en eerste term I). Daar $\|A^n\| \le \|A\|^n$ (1.76) en $\|A\| < 1$ is $\lim A^n = 0$ en dus $\lim_{n \to \infty} S_n = (I-A)^{-1}$.

1.78. $A_{n \times n} > 0 \Rightarrow |A| \le \prod_i a_{ii}$.

Bewijs met inductie:

Waar voor n = 1; stel waar voor n = n.

Partitioneer $A_{(n+1)(n+1)} = \begin{pmatrix} A_{nn} & b \\ b' & a_{n+1,n+1} \end{pmatrix}$.

Dus (zie 1.35)

$|A_{(n+1)(n+1)}| = |A_{nn}|\,|a_{n+1,n+1} - b'A_{nn}^{-1}b| \le |A_{nn}|a_{n+1,n+1} \le \prod_i^{n+1} a_{ii}$ nl.

$b'A^{-1}b \ge o$ en $a_{n+1,n+1} > o$ (zie 1.67). Hieruit volgt de

1.79. *Hadamard ongelijkheid:* $|A|^2 \le \prod_i \|a_{*i}\|^2$.

<u>Bewijs</u>:

$B := A'A$ dus (1.78) $|B| \le \prod b_{ii}$; maar $b_{ii} = a_{*i}^2 = \|a_{*i}\|^2$ en $|B| = |A|^2$ oftewel $|A|^2 \le \prod_i \|a_{*i}\|^2$.

In woorden meetkundig: het volume opgespannen door de kolommen van A (dus det A) is kleiner of gelijk aan het product van de lengten dezer kolommen.

1.80. Zij $A \in \mathcal{M}^{p}_{p,q} \Rightarrow \exists B \in \mathcal{M}^{p}_{q,p}$ met $AB = I_p$; B heet *rechterinverse* van A.
Neem $B = A'(AA')^{-1}$. Analoog:
Zij $A \in \mathcal{M}^{q}_{p,q} \Rightarrow \exists B \in \mathcal{M}^{q}_{q,p}$ met $BA = I_q$; B heet *linkerinverse* van A.
Neem $B = (A'A)^{-1}A'$.

1.81. *Hoofdwaardendécompositie*: Zij $r(A_{p\times q}) = r$. Dan bestaan er matrices $W_{pr}$, $V_{qr}$ met $W^{\underline{2}} = V^{\underline{2}} = I$ en een diagonaalmatrix $\Lambda$ zodat $A = W\Lambda^{\frac{1}{2}}V'$.

Bewijs:
Neem voor de kolommen van V orthonormale eigenvectoren behorende bij de r eigenwaarden > 0 van de matrix $A^{\underline{2}}$. Dus $A^{\underline{2}} = V\Lambda V'$. Definieer $W := AV\Lambda^{-\frac{1}{2}}$. Dan is $W^{\underline{2}} = \Lambda^{-\frac{1}{2}}V'A^{\underline{2}}V\Lambda^{-\frac{1}{2}} = I$. En $(A-W\Lambda^{\frac{1}{2}}V')^{\underline{2}} = [A(I-VV')]^{\underline{2}} = (I-VV')A^{\underline{2}}(I-VV') = 0$, dus $A = W\Lambda^{\frac{1}{2}}V'$. q.e.d.

Opmerking:
1. $\Lambda$ bevat de positieve eigenwaarden van $A^{\underline{2}}$, deze worden de *hoofdwaarden* van A genoemd (Eng. singular values).
2. De kolommen van W zijn de orthonormale eigenvectoren behorende bij de positieve eigenwaarden van AA' (zie ook 1.72) nl.:

$$AA'W = (W\Lambda^{\frac{1}{2}}V')(V\Lambda^{\frac{1}{2}}W')W = W\Lambda.$$

3. Men kan W en V aanvullen tot orthogonale matrices door toe te voegen orthonormale bases van N(A') resp. N(A). $\Lambda$ wordt dan met nullen aangevuld tot $\Lambda_{pq}$.
4. Andere versie: er bestaan reguliere matrices R en Q zodat

$$RAQ = \begin{pmatrix} I_{rr} & 0 \\ 0 & 0 \end{pmatrix}$$
. Dit volgt direct uit opmerking 3.

Dan is $R = W'$ en $Q = V\begin{pmatrix} \Lambda^{-\frac{1}{2}} & 0 \\ 0 & I \end{pmatrix}$.

1.82. A,B symmetrisch: $|I-tA||I-tB| = |I-tA-tB| \Longleftrightarrow AB = 0$.

Bewijs $\Leftarrow$ : triviaal.
Bewijs $\Rightarrow$ : $|I-tB| = |I-tA|^{-1}|I-tA-tB| = |I-tKB|$ met $K := (I-tA)^{-1}$.

<u>Opmerking</u>: I-tA is regulier als 1/t geen eigenwaarde is van A.
Dus ook is $|I-t^2B^2| = |I-t^2(KB)^2|$ oftewel sp $B^2$ = sp$(KB)^2$ =
= sp $[(I+tA+t^2A^2+ \ldots)B]^2$ = sp$[B^2+t(\ldots)+t^2(ABAB+A^2B^2+BA^2B) + O(t^3)]$

$\therefore$ sp[BABA+2(BA)(BA)'] = o. Stel BA = D. Nu is (1.75)
sp(DD+DD') $\geq$ o dus sp(DD') = o oftewel D = BA = AB = 0 (1.58).

1.83. A,B symmetrisch; AB = BA $\Longleftrightarrow$ $\exists$ orthog. H zodat $H'AH = \Lambda_1$ diag.
$H'BH = \Lambda_2$ diag.

<u>Bewijs</u> $\Leftarrow$ : $AB = H\Lambda_1H'H\Lambda_2H' = H\Lambda_1\Lambda_2H' = H\Lambda_2\Lambda_1H' = H\Lambda_2H'H\Lambda_1H' = BA$.

<u>Bewijs</u> $\Rightarrow$ :
Stel $R'AR = \Lambda_1 = \begin{pmatrix} \lambda_1 I & & 0 \\ & \ddots & \\ 0 & & \lambda_r I \end{pmatrix}$ R orthogonaal, $\lambda_i \neq \lambda_j$.

Stel R'BR =: C (dus symmetrisch)
AB = BA geeft $\Lambda_1 C = C\Lambda_1$ of uitgeschreven, C gepartitioneerd:

$$\begin{pmatrix} \lambda_1 I & & 0 \\ & \ddots & \\ 0 & & \lambda_r I \end{pmatrix} \begin{pmatrix} C_{11} & \cdots & C_{1r} \\ \vdots & & \vdots \\ C_{r1} & \cdots & C_{rr} \end{pmatrix} = \begin{pmatrix} C_{11} & \cdots & C_{1r} \\ \vdots & & \vdots \\ C_{r1} & \cdots & C_{rr} \end{pmatrix} \begin{pmatrix} \lambda_1 I & & 0 \\ & \ddots & \\ 0 & & \lambda_r I \end{pmatrix} .$$

Dit geeft $\lambda_i C_{ij} = \lambda_j C_{ij} (= C_{ji})$ en daar $\lambda_i \neq \lambda_j$ is dus $C_{ij} = 0$ voor $i \neq j$.
Stel $Q_i' C_{ii} Q_i = D_i$(diag) en $Q_i$ orthog. (dit kan daar C, dus $C_{ii}$ symmetrisch is)

$$Q := \begin{pmatrix} Q_1 & & 0 \\ & \ddots & \\ 0 & & Q_r \end{pmatrix} \quad \text{met } Q'Q = I .$$

$$\text{Nu is } Q'CQ = \begin{pmatrix} D_1 & & 0 \\ & \ddots & \\ 0 & & D_r \end{pmatrix} =: \Lambda_2 \text{ (diag.)}$$

$$\text{en } Q'\Lambda_1 Q = \begin{pmatrix} Q_1' & & 0 \\ & \ddots & \\ 0 & & Q_r' \end{pmatrix} \begin{pmatrix} \lambda_1 I & & 0 \\ & \ddots & \\ 0 & & \lambda_r I \end{pmatrix} \begin{pmatrix} Q_1 & & 0 \\ & \ddots & \\ 0 & & Q_r \end{pmatrix} = \Lambda_1 .$$

Dus $Q'R'ARQ = \Lambda_1$ èn $Q'R'BRQ = \Lambda_2$ ; noem RQ = H. q.e.d.
Gevolg: A,B > 0 èn AB symmetrisch $\Rightarrow$ AB > 0.

Idempotente matrices en projecties

Nu volgen enkele eigenschappen betreffende idempotente matrices. Deze spelen een belangrijke rol o.a. bij het opsplitsen van kwadraatsommen en het bepalen van kansdichtheden van kwadratische vormen.

1.90. P orthogonale projector $\iff P = P\underline{^2}$ (dus symmetrisch en idempotent).

Bewijs:
Zij $R^n$ Euclidische vectorruimte met orthonormale basis.

$$[\forall x,y\ (y-Py)'Px = y'(I-P')Px = o] \iff (I-P')P = 0 \iff P = P'P = P\underline{^2}\ .$$

Opmerkingen:

i) Is P een orthogonale projector op een lineaire deelruimte D dan is $P^{\perp} := I-P$ de orthogonale projector op het orthoplement $D^{\perp}$ van D, nl. I-P is weer symmetrisch en idempotent en $P^{\perp}P = 0$.

ii) Met $P_A$ zal voortaan bedoeld zijn een orthogonale projector op de deelruimte A of <A>.

iii) Is $A \perp B$ dan is $P_AP_B = 0$ nl. $(P_Ax, P_By) = x'P_AP_By = o$ voor alle x,y dus $P_AP_B = 0$ (zie 1.57).

1.91. Gevolg:

i) $P_{n\times n} = \{\frac{1}{n}\} = P_u = \frac{1}{n} U = \frac{1}{n} uu'$ is een orthogonale projector op <u> (ga na).

ii) $P_X := X(X'X)^{-1}X'$ is de orthogonale projector op <X> (ga na).
$P_X^{\perp} = I-P_X$ is de orthogonale projector op $<X>^{\perp}$ .

iii) De orthogonale projectie van een vector y op een vector x is $\frac{y'x}{x'x}\,x$.
Is $x = u$, $y' = (y_1,\dots,y_n)$ dan wordt dit juist
$\bar{y} := (\bar{y}_{.},\dots,\bar{y}_{.})'$ met $\bar{y}_{.} := \Sigma y_i/n$.

1.92. $<P_X> = <X>$ nl. $P_XX = X$ dus $<X> \subset <P_X>$ (zie 1.48).
Verder is $r(P_X) = \dim<X> = r(X)$ dus $<X> = <P_X>$.

1.93. $A_{p\times p}$ niet defect en $\forall i\ \lambda_i = o$ of $1 \iff A$ idempotent.

<u>Bewijs</u> $\Rightarrow$ : $A = B\Lambda B^{-1}$ geeft $A^2 = B\Lambda B^{-1}B\Lambda B^{-1} = B\Lambda^2 B^{-1} = B\Lambda B^{-1} = A$ daar $\Lambda^2 = \Lambda$.

$\Leftarrow$ $A^2 = A$; $Ax = \lambda x$; $A^2x = \lambda^2 x = \lambda x$; $\lambda(\lambda-1)x = o$, $\lambda = o$ of $1$. Dus $\langle A\rangle = E_{\lambda=1}$. Stel $r(A) = r$, dan $\dim N(A) = p-r$. Vul $B$ aan met $p-r$ onafhankelijke eigenvectoren uit $N(A)$ bij $\lambda = o$. Dus $A$ niet defect!

<u>Opmerking</u>: Is $A$ symmetrisch, dan is $A$ s.p.d. tenzij $A = I_p$.

1.94. $A$ idempotent: $sp(A) = r(A)$ nl. (met 1.93)
$sp\ A = sp(B\Lambda B^{-1}) = sp\ \Lambda = r(A)$. De relatie is zeer belangrijk!

1.95. $P_D$ orthogonale projector op r-dim.deelruimte $D \iff P = G^{\underline{2}}$ met $GG' = I_r$.

<u>Bewijs</u> $\Rightarrow$ : $P = H\Lambda H' = H\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}H' = H\binom{I}{0}(I\quad 0)H' =: G'G = G^{\underline{2}}$ (zie ook 1.64).

$\Leftarrow$ $G^{\underline{2}}$ is volgens 1.90 een orthogonale projector! En met 1.92 is $\dim D = r(P_X) = r(G^{\underline{2}}) = r(G) = r(GG') = r(I_r) = r$ (met 1.45).

<u>Opmerking</u>: Daar $PG' = G'GG' = G'$ zijn de kolommen van $G'$ orthonormale eigenvectoren bij $\lambda = 1$.

1.96. $\Sigma A\Sigma'$ orthog projector $\Leftrightarrow \Sigma^{\underline{2}}A\Sigma^{\underline{2}}A\Sigma^{\underline{2}} = \Sigma^{\underline{2}}A\Sigma^{\underline{2}}$ en $\Sigma^{\underline{2}}A\Sigma^{\underline{2}}$ symmetrisch

$1^e$ $\quad \Sigma A\Sigma' = \Sigma A'\Sigma' \Leftrightarrow \Sigma^{\underline{2}}A\Sigma^{\underline{2}} = \Sigma^{\underline{2}}A'\Sigma^{\underline{2}}$ (met 1.60).

$2^e$ $\quad \Sigma A\Sigma'\Sigma A\Sigma' = \Sigma A\Sigma' \Leftrightarrow \Sigma^{\underline{2}}A\Sigma^{\underline{2}}A\Sigma^{\underline{2}} = \Sigma^{\underline{2}}A\Sigma^{\underline{2}}$ (met 1.60).

1.97. Zij $B$ deelruimte $\subset V$; $a \in V$ ; dan geldt $d[a,B] = \|a-P_B a\|$ nl
$\forall b \in B$ is $\|a-b\|^2 = \|a-P_B a\|^2 + \|P_B a-b\|^2$ dus $\|a-P_B a\|^2 \le \|a-b\|^2$ (zie 1.9).

1.98. <u>Probleem</u>. Bepaal bij gegeven matrix $Y_{n\times p}$ en gegeven $k < n$, een matrix $X_{n\times k}$ van volle rang en een matrix $B$, zodat $\|Y-XB\|$ minimaal is.

<u>Oplossing</u>. De matrices van de orde $n \times p$ vormen een lineaire vectorruimte en met het inproduct als gedefinieerd in 1.8 een Euclidische vectorruimte. Deze wordt metrisch met $d[A,B] := \|A-B\|$ als afstandsfunctie. We proberen dus $Y$ "zo goed mogelijk" door $XB$ te benaderen.

Definieer P als de orthogonale projector op de, nog onbekende, k-dimensionale lineaire deelruimte $\langle X\rangle$ opgespannen door de kolommen van X; $P^{\perp}$ is de orthogonale projectie op het orthoplement $\langle X\rangle^{\perp}$ van $\langle X\rangle$. Uit 1.32 en de projectiestelling volgt nu

$$f(X,B) := \|Y-XB\|^2 = \sum_j \|y_{*j}-Xb_{*j}\|^2 = \sum_j \|P(y_{*j}-Xb_{*j})\|^2 + \sum_j \|P^{\perp}(y_{*j}-Xb_{*j})\|^2$$

oftewel

$$f(X,B) = \|P(Y-XB)\|^2 + \|P^{\perp}(Y-XB)\|^2$$

$$f(X,B(X)) := \min_B f(X,B) = \|P^{\perp}Y\|^2$$

voor $(Y-XB)'X = 0$, $B(X) = (X'X)^{-1}X'Y$. Nu is $YY' = H\Lambda H'$ (H orthonormale eigenvectoren; $\Lambda$ eigenwaarden met $\lambda_1 \geq \ldots \geq \lambda_n$). Dus

$$\|P^{\perp}Y\|^2 = sp(P^{\perp}YY') = sp(P^{\perp}H\Lambda H') = \|P^{\perp}H\Lambda^{\frac{1}{2}}\|^2 = \sum_1^n \lambda_i \|P^{\perp}h_{*i}\|^2 =: \sum_1^n \lambda_i p_i$$

met $0 \leq p_i \leq 1$ en $\sum_1^n p_i = sp(P^{\perp}HH') = sp\, P^{\perp} = r(P^{\perp}) = n-k$ (zie 1.94).

Definieer $s_i = \sum_1^i p_j$ voor $i = 1,\ldots,n$; $s_0 = \lambda_{n+1} = 0$. Dan is

$$\sum_1^n \lambda_i p_i = \sum_1^n \lambda_i(s_i-s_{i-1}) = \sum_1^n (\lambda_i-\lambda_{i+1})s_i \,.$$

Nu is $s_n = n-k$, dus $s_i \geq i-k$ daar $p_i \leq 1$. Verder is $s_i \geq 0$, dus

$$\sum_1^n (\lambda_i-\lambda_{i+1})s_i \geq \sum_{k+1}^n (\lambda_i-\lambda_{i+1})(i-k) = \sum_{k+1}^n \lambda_i \,,$$

oftewel

$$f(X_0,B_0) := \min_X f(X,B(X)) = \min_X \sum_1^n \lambda_i \|P^{\perp}h_{*i}\|^2 = \sum_{k+1}^n \lambda_i$$

voor $\|P^{\perp}h_{*i}\| = 0$, $i = 1,\ldots,k$, en $\|P^{\perp}h_{*i}\| = 1$ voor $i = k+1,\ldots,n$.

<u>Conclusie</u>: Kies voor de kolommen van $X_0$ orthon. eigenvect. van YY' behorende bij de k grootste eigenwaarden. Inderdaad is dan $\|P^{\perp}h_{*i}\| = \|h_{*i}\| = 1$ voor $h_{*i} \in \langle X_0\rangle^{\perp}$ en $\|P^{\perp}h_{*i}\| = 0$ voor $h_{*i} \in \langle X_0\rangle$.

Opmerking. We zien dat $XB = X(X'X)^{-1}X'Y = PY$. Ook zien we dat, wil deze benadering van Y door XB zin hebben, $r(YY') = r(Y) = r(\Lambda) > k$ moet zijn (dus $p > k$).

Uit dit probleem volgen direct enkele stellingen betreffende extrema:
Onderstel $A_{n\times n} \geq 0$ met eigenwaarden $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$; $r(X_{n\times k}) = k$.
Dan is

1.99. $\min\limits_X sp(P_X A) = \sum\limits_{n-k+1}^{n} \lambda_i$ = som van de k kleinste eigenwaarden van A.

De kolommen van X zijn de bijbehorende orthonormale eigenvectoren.
Analoog is natuurlijk:

$$\max_X sp(P_X A) = \sum_1^k \lambda_i = \text{som van de k grootste eigenwaarden van A, nl.}$$

$$= \max_X sp[(I-P_X^{\perp})A] = \sum_1^n \lambda_i - \min_X sp(P_X^{\perp}A) = \sum_1^n \lambda_i - \sum_{k+1}^{n} \lambda_i = \sum_1^k \lambda_i .$$

Bijzondere gevallen hiervan zijn:

1.100. $\max\limits_{X^2=I} sp(X'AX) = \sum\limits_1^k \lambda_i$ ,

nl. $sp(X'AX) = sp(XX'A) = sp(P_X A)$, nl. $P_X = XX'$ is symmetrisch en idempotent.
Voor X = x wordt dit:

$\max\limits_{|x|=1} x'Ax = \lambda_1$ = grootste eigenwaarde van A, en x bijbehorende eigenvector.

Dit is natuurlijk hetzelfde als $\max\limits_{x\neq 0} \dfrac{x'Ax}{x'x} = \lambda_1$ (geldt ook als A symmetrisch is).

Opmerking:

1) het quotiënt $q(x) := \dfrac{x'Ax}{x'x}$ wordt het Rayleigh quotiënt genoemd.

2) $\max\limits_x \dfrac{x'Ax}{x'Bx} = \lambda_1$ = grootste eigenwaarde van $B^{-1}A$.

   Hierbij is A symmetrisch, $B > 0$ (ga na).

## § 2. Enkele begrippen uit de mathematische statistiek

In deze paragraaf worden enkele begrippen uit de mathematische statistiek nog eens herhaald en enkele nieuwe begrippen ingevoerd, o.a. de maximum likelihood methode die veel wordt toegepast om schatters te vinden, en de variantie-covariantiematrix die een centrale rol speelt.

### a. Definities, notaties

2.1. De $\mathcal{E}$-*operator*

$$\mathcal{E}\underline{x} := \sum p_i x_i \quad , \ \underline{x} \text{ diskreet, mits de som absoluut konvergeert,}$$

en

$$\mathcal{E}\underline{x} := \int_{-\infty}^{\infty} x f(x)dx \ , \ \underline{x} \text{ kontinu, mits de integraal absoluut konvergeert.}$$

Stelling. $\mathcal{E}_x g(\underline{x}) := \int_{-\infty}^{\infty} g(x)f(x)dx = \mathcal{E}g(\underline{x})$ .

Ook meerdimensionaal:

$$\mathcal{E}g(\underline{x}_1,\ldots,\underline{x}_n) = \int_G g(x_1,\ldots,x_n)f(x_1,\ldots,x_n)dx_1 \ldots dx_n \ .$$

$\mathcal{E}g(\underline{x})$ heet de *verwachting* van $g(\underline{x})$.

Opmerking. Onderstreepte letters zijn stochastische grootheden.

2.2. Een *stochastische* matrix $\underline{A} = \{\underline{a}_{ij}\}$ is een matrix waarvan de elementen $\underline{a}_{ij}$ stochastisch zijn (niet te verwarren met een Markofmatrix).

2.3. De *verwachting* van $\underline{A}$ is $\mathcal{E}\underline{A} := \{\mathcal{E}\underline{a}_{ij}\}$.

2.4. $\underline{x}$ is *isomoor* met $\underline{y}$ (notatie $\underline{x} \simeq \underline{y}$) betekent $\underline{x}$ en $\underline{y}$ hebben dezelfde kansverdeling; $\underline{x} \sim N(\mu,\sigma^2)$ betekent $\underline{x}$ is normaal verdeeld met gemiddelde $\mu$ en variantie $\sigma^2$.

2.5. *Karakteristieke funktie* $\varphi_{\underline{x}}(t) := \mathcal{E}e^{it\underline{x}}$ (t reëel). Analoog indien $\underline{x}$ een p-dimensionale variabele (vektor) is: $\varphi_{\underline{x}}(t) := \mathcal{E}e^{it'\underline{x}}$ met $t' = (t_1,\ldots,t_p)$ reëel.

2.6. Een *aselecte steekproef* ter grootte n van de p-dimensionale variabele $\underline{x} = (\underline{x}_1,\ldots,\underline{x}_p)'$ is een matrix $\underline{X}_{n\times p}$ van n onafhankelijke, met $\underline{x}$ isomore stochastische vektoren. De rijen geven dus de herhalingen aan.

2.7. Een *schatter* van een populatiegrootheid (parameter) op basis van een steekproef $\underline{X}_{n\times p}$ is een funktie van deze $\underline{X}_{n\times p}$.

2.8. Een schatter $\underline{t}$ voor de parameter $\vartheta$ heet *zuiver* als geldt:

$$\mathcal{E}\underline{t} = \vartheta \quad \text{(voor alle } \vartheta \in \text{parameterruimte).}$$

2.9. Een *schatting* is een realisatie van een schatter, dus t is een realisatie van $\underline{t}$.

2.10. De *aannemelijkheidsfunktie (likelihood function)* van een onafhankelijke steekproef $\underline{X}_{n\times p}$ ter grootte n uit de verdeling $f(x;\vartheta)$ waarin $\underline{x} = (\underline{x}_1,\ldots,\underline{x}_p)'$ en $\vartheta = (\vartheta_1,\ldots,\vartheta_k)'$, is de simultane kansdichtheid van $\underline{X}_{n\times p}$ opgevat als funktie van $\vartheta$:

$$\underline{L}(\vartheta) = \prod_1^n f(x_{i*};\vartheta) \qquad \text{waarin} \quad \underline{x}_{i*} = (\underline{x}_{i1},\ldots,\underline{x}_{ip})' \ .$$

2.11. De *meest aannemelijke schatter* voor $\vartheta$ is die $\underline{\hat{\vartheta}}$ die $\underline{L}(\vartheta)$ maximaliseert.

2.12. De *likelihood-ratio toets* voor de nulhypothese $H_0 : \vartheta \in \omega$ tegen de alternatieve $H_1 : \vartheta \notin \omega$, heeft als kritieke zone die steekproefuitkomsten waarvoor

$$0 < \lambda = \frac{L_\omega}{L_\Omega} := \frac{\sup_{\vartheta\in\omega} L(\vartheta)}{\sup_{\vartheta\in\Omega} L(\vartheta)} < \lambda_\alpha \qquad (\Omega = \text{parameterruimte}).$$

2.13. *Populatie* (ééndimensionaal)
*populatiegrootheid* (parameter)

*pop.gemiddelde* $\mu := \mathcal{E}\underline{x}$

*pop.variantie* $\sigma^2 := \mathcal{E}(\underline{x} - \mu)^2$

*pop.spreiding* $\sigma := \sqrt{\sigma^2}$

*Steekproef* $(\underline{x}_1,\ldots,\underline{x}_n)$
*steekproefgrootheid* (schatter)

*steekpr.gemiddelde* $\underline{\bar{x}} := \sum_1^n \underline{x}_i/n$

*steekpr.variantie* $\underline{s}^2 := \sum(\underline{x}_i-\underline{\bar{x}})^2/(n-1)$

*steekpr.spreiding* $\underline{s} := \sqrt{\underline{s}^2}$ .

| *Populatie* (tweedimensionaal) | *steekproef* $\underline{x}_1,\underline{y}_1;\ldots;\underline{x}_n,\underline{y}_n$ |
|---|---|
| *pop.covariantie* $\mathrm{cov}(\underline{x},\underline{y}) =$ | *steekpr.covariantie* $s^2(\underline{x},\underline{y}) :=$ |
| $= \sigma^2(\underline{x},\underline{y}) := \mathcal{E}[(\underline{x}-\mu_x)(\underline{y}-\mu_y)]$ | $= \sum_1^n (\underline{x}_i - \bar{\underline{x}})(\underline{y}_i - \bar{\underline{y}})/(n-1)$ |
| *pop.correlatiecoëfficiënt* | *steekpr.correlatiecoëfficiënt* |
| $\rho(\underline{x},\underline{y}) = \dfrac{\sigma^2(\underline{x},\underline{y})}{\sigma(\underline{x})\sigma(\underline{y})}$ | $r(\underline{x},\underline{y}) = \dfrac{s^2(\underline{x},\underline{y})}{s(\underline{x})s(\underline{y})}$ . |

2.14. De *variantie-covariantiematrix* van de stochastische vektor $\underline{x} = (\underline{x}_1,\ldots,\underline{x}_p)'$ is

$$\mathrm{VAR}(\underline{x}) = \Sigma^2(\underline{x}) := \mathcal{E}(\underline{x}-\mu)(\underline{x}-\mu)' = \Sigma^2 \quad \text{met } \mu = \mathcal{E}\underline{x} = (\mu_1,\ldots,\mu_p)' \text{ en } \mu_i = \mathcal{E}\underline{x}_i \ .$$

Inderdaad is

$$(\Sigma^2)_{ij} = \mathcal{E}(\underline{x}_i - \mu_i)(\underline{x}_j - \mu_j) = \sigma^2_{ij} \ .$$

Zijn de $\underline{x}_i$ alle gestandaardiseerd, dan wordt $\Sigma^2 = \mathcal{R}$ de *correlatiematrix*.

Opmerking. In de literatuur is $\Sigma$ gebruikelijk i.p.v. $\Sigma^2$. Toch biedt deze laatste notatie voordelen, o.a. door de analogie met $\sigma^2$ voor $p = 1$. $\Sigma^2 = \Sigma'\Sigma$. Dat dit geoorloofd is, dus m.a.w. dat $\mathrm{VAR}(\underline{x}) \geq 0$ is, wordt bewezen in 2.39. Ook is $\Sigma^2 = \Sigma\Sigma$, nu met symmetrische $\Sigma$ (zie 1.64).

2.15. De *covariantiematrix* der vektoren $\underline{x} = (\underline{x}_1,\ldots,\underline{x}_p)'$ en $\underline{y} = (\underline{y}_1,\ldots,\underline{y}_q)'$ is

$$\mathrm{COV}(\underline{x},\underline{y}) := \mathcal{E}(\underline{x}-\mu_x)(\underline{y}-\mu_y)' \ .$$

In het bijzonder is $\mathrm{COV}(\underline{x},\underline{x}) = \mathrm{VAR}\ \underline{x}$ en $\mathrm{COV}(\underline{x},\underline{y}) = \mathrm{COV}'(\underline{y},\underline{x})$.

2.16. Een lineaire funktie $a'\underline{y}$ is een BLUE (*best linear unbiased estimator*) voor $c'\beta$ als $\mathcal{E}(a'\underline{y}) = c'\beta$ en $\mathrm{var}(a'\underline{y})$ minimaal is (onder alle zuivere lineaire schatters).

Opmerking. Ook gebruikelijke definitie voor BLUE is: $\underline{t} = (\underline{t}_1,\ldots,\underline{t}_p)'$ is een BLUE voor $\vartheta$ als $\mathcal{E}\underline{t} = \vartheta$, $\underline{t}$ lineair en $\underline{t}$ minimale gegeneraliseerde variantie heeft (binnen de klasse van alle lineaire zuivere schatters).

2.17. $|\Sigma^2|$ wordt de *gegeneraliseerde variantie* van een verdeling genoemd.

b. Eigenschappen

2.30. $\sigma^2(\underline{x}) = \mathcal{E}\underline{x}^2 - (\mathcal{E}\underline{x})^2$; $\sigma^2(\underline{x},\underline{y}) = \mathcal{E}(\underline{x}\underline{y}) - \mathcal{E}\underline{x}\,\mathcal{E}\underline{y}$.

2.31. Zijn $\underline{x}$ en $\underline{y}$ stochastisch onafhankelijk, dan geldt $\mathcal{E}(\underline{x}\underline{y}) = \mathcal{E}\underline{x}\,\mathcal{E}\underline{y}$ en dus zijn $\underline{x}$ en $\underline{y}$ *ongecorreleerd* oftewel $\rho(\underline{x},\underline{y}) = 0$ (niet omgekeerd).

2.32. $\mathrm{var}(\alpha\underline{x} + \beta\underline{y}) = \alpha^2\,\mathrm{var}\,\underline{x} + \beta^2\,\mathrm{var}\,\underline{y} + 2\alpha\beta\,\mathrm{cov}(\underline{x},\underline{y})$.

2.33. Is $\underline{y} = (\underline{x} - \mu)/\sigma$, dan is $\mu(\underline{y}) = 0$ en $\sigma(\underline{y}) = 1$; $\underline{y}$ heet *gestandaardiseerd*. Analoog geldt: als $y_i = (x_i - \bar{x})/s(x)$, dan is $\bar{y} = 0$ en $s(y) = 1$.

2.34. Is $\underline{u} = \alpha\underline{x} + \beta$ en $\underline{v} = \gamma\underline{y} + \delta$ met $\alpha\gamma \neq 0$, dan is $\rho^2(\underline{u},\underline{v}) = \rho^2(\underline{x},\underline{y})$, m.a.w. $\rho^2$ is onder deze transformaties invariant.

2.35. $\mathcal{E}(A\underline{B} + \underline{C}) = A\,\mathcal{E}\underline{B} + \mathcal{E}\underline{C}$.

Bewijs:

$$[\mathcal{E}(A\underline{B} + \underline{C})]_{ij} = \mathcal{E}(\sum_k a_{ik}\underline{b}_{kj} + \underline{c}_{ij}) = \sum_k a_{ik}\,\mathcal{E}\underline{b}_{kj} + \mathcal{E}\underline{c}_{ij} =$$

$$= \sum_k a_{ik}(\mathcal{E}\underline{B})_{kj} + (\mathcal{E}\underline{C})_{ij} = (A\mathcal{E}\underline{B})_{ij} + (\mathcal{E}\underline{C})_{ij} = (A\mathcal{E}\underline{B} + \mathcal{E}\underline{C})_{ij}\ .$$

2.36. De *verwachting van een kwadratische vorm*:

$$\mathcal{E}(\underline{x}'A\underline{x}) = \mathrm{sp}(A\Sigma^2) + \mu'A\mu\ .\ \text{Analoog}\ \mathcal{E}(\underline{x}'A\underline{y}) = \mathrm{sp}(A\Sigma_{21}) + \mu_x'A\mu_y\ .$$

Bewijs: $\Sigma^2 = \mathcal{E}(\underline{x} - \mu)(\underline{x} - \mu)' = \mathcal{E}\underline{x}\underline{x}' - \mu\mu'$

dus $A\Sigma^2 = \mathcal{E}(A\underline{x}\underline{x}') - A\mu\mu'$

en $\mathrm{sp}(A\Sigma^2) = \mathcal{E}(\mathrm{sp}\,A\underline{x}\underline{x}') - \mathrm{sp}(A\mu\mu') = \mathcal{E}(\underline{x}'A\underline{x}) - \mu'A\mu$ (zie 1.32).

Is $\Sigma^2 = \sigma^2 I$ en $A = P$ een orthogonale projectie met $r(P) = \nu$ dan ontstaat:

$$\mathcal{E}\|P\underline{x}\|^2 = \mathcal{E}(\underline{x}'P\underline{x}) = \mathrm{sp}(\sigma^2 P) + \mu'P\mu = \nu\sigma^2 + \mu'P\mu.$$

2.37. $\mathrm{COV}(A\underline{x} + B\underline{y}, C\underline{z}) = A\,\mathrm{COV}(\underline{x},\underline{z})C' + B\,\mathrm{COV}(\underline{y},\underline{z})C'$.

Bewijs:

$$\mathrm{COV}(A\underline{x} + B\underline{y}, C\underline{z}) = \mathcal{E}[(A\underline{x} + B\underline{y} - A\mu_x - B\mu_y)(C\underline{z} - C\mu_z)'] =$$

$$= \mathcal{E}[\{A(\underline{x} - \mu_x) + B(\underline{y} - \mu_y)\}(\underline{z} - \mu_z)'C'] = \mathcal{E}[A(\underline{x} - \mu_x)(\underline{z} - \mu_z)'C'] +$$

$$+ \mathcal{E}[(B\underline{y} - \mu_y)(\underline{z} - \mu_z)'C'] = A\,\mathrm{COV}(\underline{x},\underline{z})C' + B\,\mathrm{COV}(\underline{y},\underline{z})C' \ .$$

2.38. Gevolg: $\mathrm{VAR}(A\underline{x}) = A\Sigma^2 A'$, nl. met $B = 0$, $A = C$ en $\underline{z} = \underline{x}$.
In het bijzonder als $\underline{y} = A\underline{x}$, $\Sigma^2(\underline{x}) = \sigma^2 I$ en $A$ orthogonaal geldt:

$$\Sigma^2(\underline{y}) = A\Sigma^2(\underline{x})A' = A\sigma^2 IA' = \sigma^2 AA' = \Sigma^2(\underline{x}) \ .$$

2.39. $0 < \mathrm{var}(\alpha'\underline{x}) = \alpha'\Sigma^2\alpha$, d.w.z. $\Sigma^{\underline{2}} > 0$ (tenzij $\alpha'\underline{x}$ constant is oftewel de $\underline{x}_i$ lineair afhankelijk).
Hieruit blijkt dus dat de notatie $\Sigma^{\underline{2}}$ voor VAR $\underline{x}$ geoorloofd is, daar $\Sigma^{\underline{2}} := \Sigma'\Sigma$ dus (s.) p.d. is.

2.40. Is $\hat{\underline{\vartheta}}$ een maximum likelihoodschatter voor $\vartheta$ en f een één-éénduidige funktie, dan is $f(\hat{\underline{\vartheta}})$ een maximum likelihoodschatter voor $f(\vartheta)$. (Zie [1] pag.48.)

2.41. Als $\varphi_{\underline{x}}(t) = \varphi_{\underline{y}}(t)$ voor alle t, dan is $x \approx y$ (zie Cramèr, pag. 94).

2.42. $\varphi_{A\underline{x}+\beta}(t) = \mathcal{E}e^{it'(A\underline{x}+\beta)} = e^{it'\beta}\,\varphi_{\underline{x}}(A't)$, i.h.b. voor $A = \alpha$:
$\varphi_{\alpha\underline{x}+\beta} = e^{i\beta t}\,\varphi_{\underline{x}}(\alpha t)$.

2.43. Is $\underline{x} \sim N(0,1)$, dan is $\varphi_{\underline{x}}(t) = e^{-\frac{1}{2}t^2}$.
Bewijs:

$$\varphi_{\underline{x}}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx}\, e^{-\frac{1}{2}x^2} dx = \frac{e^{-\frac{1}{2}t^2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-it)^2} dx = e^{-\frac{1}{2}t^2} \ .$$

Is $\underline{y} \sim N(\mu,\sigma^2)$ oftewel $\underline{y} = \sigma\underline{x} + \mu$, dan is volgens 2.42

$$\varphi_{\underline{y}}(t) = e^{i\mu t}\, e^{-\frac{1}{2}(\sigma t)^2} = \exp(i\mu t - \tfrac{1}{2}\sigma^2 t^2) \ .$$

2.44. $\varphi_{\Sigma \underline{x}_i}(t) = \prod_i \varphi_{\underline{x}_i}(t)$ indien alle $\underline{x}_i$ onderling stochastisch onafhankelijk zijn.

In het bijzonder is dus $\varphi_{\underline{x}+\underline{y}}(t) = \varphi_{\underline{x}}(t)\varphi_{\underline{y}}(t)$, nl.

$$\mathcal{E}e^{it(\underline{x}+\underline{y})} = \iint e^{it(x+y)} f(x)g(y)dxdy = \int e^{itx} f(x)dx \int e^{ity} g(y)dy = \varphi_{\underline{x}}(t)\varphi_{\underline{y}}(t) .$$

Dit niet te verwarren met het volgende:

Is $\underline{x} = (\underline{x}_1,\ldots,\underline{x}_p)'$, waarbij alle $\underline{x}_j$ onderling onafhankelijk zijn, dan is

$$\varphi_{\underline{x}}(t) = \prod_{j=1}^{p} \varphi_{\underline{x}_j}(t_j) \qquad \text{met } t = (t_1,\ldots,t_p)'$$

(Ook omgekeerd, zie Cramèr pag. 266,296).

2.45. Stel $\varphi_{\underline{x}}(t) = \mathcal{E}e^{it'\underline{x}} = \mathcal{E}e^{i(t_1\underline{x}_1+t_2\underline{x}_2)}$, dan is $\dfrac{\partial^2\varphi(0,0)}{\partial t_1 \partial t_2} = -\ \mathcal{E}(\underline{x}_1\underline{x}_2)$, nl.:

$$\frac{\partial\varphi}{\partial t_1} = \mathcal{E}i\underline{x}_1 e^{i(t_1\underline{x}_1+t_2\underline{x}_2)} \quad ; \quad \frac{\partial^2\varphi}{\partial t_1 \partial t_2} = -\ \mathcal{E}(\underline{x}_1\underline{x}_2)e^{i(t_1\underline{x}_1+t_2\underline{x}_2)} .$$

2.46. Is $\varphi_{\underline{x}}(t)$, $\underline{x} = (\underline{x}_1,\ldots,\underline{x}_p)'$, bekend, dan vinden we $\varphi_{\underline{x}_j}(t_j) = \mathcal{E}e^{it_j\underline{x}_j}$ uit $\varphi_{\underline{x}} = \mathcal{E}e^{it'\underline{x}}$ door $t' = (0,\ldots,t_j,0,\ldots,0)$ te nemen.

2.47. <u>Transformatie van stochastische variabelen</u>

Zij $\underline{x} = (\underline{x}_1,\ldots,\underline{x}_p)'$ een stochastische vektor met simultane kansdichtheid $f(x) = f(x_1,\ldots,x_p)$.

Beschouw de één-éénduidige transformaties $\underline{y}_i = y_i(\underline{x}_1,\ldots,\underline{x}_p)$ met resp. inversen $\underline{x}_i = x_i(\underline{y}_1,\ldots,\underline{y}_p)$, $i = 1,\ldots,p$.

Dan is de simultane kansdichtheid van $\underline{y} = (\underline{y}_1,\ldots,\underline{y}_p)'$:

$$g(y) = f[x_1(y_1,\ldots,y_p),\ldots,x_p(y_1,\ldots,y_p)] \text{ abs } \left|\frac{\partial x}{\partial y}\right| .$$

Hierin is $\left|\frac{\partial x}{\partial y}\right|$ de *determinant van Jacobi* $\left|\left\{\frac{\partial x_i}{\partial y_j}\right\}\right|$.

Is in het bijzonder $\underline{y} = A\underline{x}$ met $|A| \neq 0$, dan is $g(y) = f(A^{-1}y)$ abs $|A^{-1}|$.

Zijn x en y vectoren en is $y = y(x)$, dan is $\frac{\partial y}{\partial x} := \left\{\frac{\partial y_i}{\partial x_j}\right\}$ .

Is $y = Ax$, dan is $\frac{\partial y}{\partial x} = A$.

## § 3. De multinormale verdeling

De ééndimensionale normale verdeling treedt vaak op omdat het te bestuderen effekt veelal de som is van meerdere onafhankelijke variabelen.
Evenals de centrale limietstelling leidt tot de normale verdeling voor een enkelvoudige variabele, zo leidt ook de algemene centrale limietstelling tot de multinormale verdeling van meerdere variabelen.
Dit is een rechtvaardiging voor de centrale plaats die de multinormale verdeling in de multivariate analyse inneemt. Bovendien is het erg plezierig dat deze verdeling wiskundig goed hanteerbaar is.

Stel $\underline{y} = (\underline{y}_1,\dots,\underline{y}_p)'$ met $\underline{y}_i \sim N(0,1)$ en alle $\underline{y}_i$ zijn stochastisch onafhankelijk. De vektor $\underline{y}$ heet dan p-*dim. standaardnormaal verdeeld*: $\underline{y} \sim N_p(0,I)$.
Dan heet $\underline{x} = \mu + \Sigma'\underline{y}$ $\quad |\Sigma| > 0$ $\quad$ p-*dim. normaal verdeeld*: $\underline{x} \sim N_p(\mu,\Sigma^2)$, nl. $\mathcal{E}\underline{x} = \mu$ en VAR $\underline{x} = \Sigma' I \Sigma = \Sigma^2$; andere notatie $\underline{y} \simeq \underline{\chi}_p \sim N_p(0,I)$.

Welke *verdeling* heeft $\underline{x}$ ?

Nu is $f_y(y) = \prod_1^p f_i(y_i) = (2\pi)^{-\frac{1}{2}p}\, e^{-\frac{1}{2}y'y}$, dus volgens 2.47 is

$$f_x(x) = f_y[\Sigma^{-T}(x - \mu)] \text{ abs } |\Sigma^{-1}| = |\Sigma|^{-1}(2\pi)^{-\frac{1}{2}p} \exp[-\tfrac{1}{2}(x-\mu)'\Sigma^{-2}(x-\mu)] \;.$$

Wat is de *karakteristieke functie*?

Volgens 2.43 en 2.44 is $\varphi_{\underline{y}}(t) = \prod_1^p \varphi_{\underline{y}_j}(t_j) = \prod_j e^{-\frac{1}{2}t_j^2} = \exp(-\tfrac{1}{2}t't)$ en dus volgens 2.42 is

$$\varphi_{\underline{x}}(t) = e^{it'\mu}\, \varphi_{\underline{y}}(\Sigma t) = \exp(it'\mu - \tfrac{1}{2}t'\Sigma^2 t)$$

(vergelijk met 2.43).

Samengevat: is $\underline{x}$ p-dim normaal verdeeld met parameters $\mu$, $\Sigma^2$, dan is

3.1. $$f(x) = |\Sigma|^{-1}(2\pi)^{-\frac{1}{2}p} \exp[-\tfrac{1}{2}(x-\mu)'\Sigma^{-2}(x-\mu)]$$

en

3.2. $$\varphi_{\underline{x}}(t) = \exp(it'\mu - \tfrac{1}{2}t'\Sigma^2 t) \;.$$

Het geval p = 2, de *tweedimensionale normale verdeling*:

$$\Sigma^{-2} = \begin{pmatrix} 1/\sigma_1^2 & -\rho/\sigma_1\sigma_2 \\ -\rho/\sigma_1\sigma_2 & 1/\sigma_2^2 \end{pmatrix} \frac{1}{(1-\rho^2)}$$

en

$$f(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]\right\} .$$

Is $\rho = 0$ dan is $f(x_1,x_2) = f_1(x_1)f_2(x_2)$ dus $\underline{x}_1$ en $\underline{x}_2$ stochastisch *onafhankelijk*.

Daar $|\Sigma|^2 = \sigma_1^2\sigma_2^2(1-\rho^2) > 0$ is $|\rho| < 1$.
Doorsnijding met vlak $f(x)$ = constant geeft ellipsen, daar dan $(x-\mu)'\Sigma^{-2}(x-\mu)$ = constant en de eigenwaarden van $\Sigma^{-2}$ alle positief zijn (zie 1.63). Het middelpunt is $\mu = (\mu_1,\mu_2)'$.

3.3. <u>Theorema</u>:

Is $\underline{x} \sim N_p(\mu,\Sigma^2)$ en $A_{m\times p}$ heeft rang m, dan is

$$\underline{y} = \alpha + A\underline{x} \sim N_m(\alpha + A\mu, A\Sigma^2 A') .$$

<u>Bewijs</u>: Dit volgt direkt uit 2.42, 3.2 en 2.41, nl.:

$$\varphi_{\underline{y}}(t) = e^{it'\alpha}\,\varphi_{\underline{x}}(A't) = e^{it'\alpha}\exp(it'A\mu - \tfrac{1}{2}t'A\Sigma^2A't) =$$

$$= \exp[it'(\alpha + A\mu) - \tfrac{1}{2}t'(A\Sigma^2A')t]$$

en daar volgens 2.41 de karakteristieke funktie éénduidig de verdelingsdichtheid bepaalt, is het gestelde bewezen. Volgens 1.66 is $A\Sigma^2A' > 0$.

<u>Enkele belangrijke konsekwenties</u>

3.4. Is $\underline{x} \sim N_p(0,\sigma^2 I)$ en $A_{r\times p}$ zodat $AA' = I$, dan is

$$\underline{y} = A\underline{x} \sim N_r(0,\sigma^2 I) .$$

Is p = r (dus A orthogonaal), dan is dus $\underline{y} = A\underline{x} \simeq \underline{x}$.

3.5. Neem $A = (0,\dots,1,0,\dots,0)$ dan volgt direkt $A\underline{x} = \underline{x}_i \sim N(\mu_i,\sigma^2_{ii})$, oftewel elke komponent van $\underline{x}$ is ook normaal verdeeld.
Pas op: het omgekeerde is niet waar: Als van een vektor $\underline{x}$ elke komponent normaal verdeeld is, hoeft $\underline{x}$ nog niet normaal verdeeld te zijn (zie appendix A).

3.6. Neem $A = \alpha' = (\alpha_1,\dots,\alpha_p)$, dan volgt uit dit theorema

$$\underline{y} = \alpha'\underline{x} = \alpha_1\underline{x}_1 + \dots + \alpha_p\underline{x}_p \sim N(\alpha'\mu,\alpha'\Sigma^2\alpha)$$

oftewel in woorden: Een lineaire combinatie van normaal verdeelde variabelen die een multinormale verdeling volgen, is weer normaal verdeeld.
Pas op: een lineaire combinatie van normaal verdeelde variabelen zonder meer hoeft niet normaal verdeeld te zijn (zie appendix A).

In het bijzonder geeft $\alpha' = \frac{1}{n}(1,\dots,1)$ en $\underline{x} \sim N_n(\mu,\sigma^2 I)$ met $\mu = (\mu,\dots,\mu)'$
$\alpha'\underline{x} = \bar{\underline{x}} \sim N(\mu,\sigma^2/n)$.

Soms definieert men een multinormaal verdeelde variabele met behulp van

3.7. Theorema

Als elke lineaire kombinatie van de komponenten van een vektor $\underline{x}$ normaal verdeeld is, dan is $\underline{x}$ multinormaal verdeeld (dichtheidsvrije definitie).

Bewijs: Stel $\underline{x} \sim F_p(\mu,\Sigma^2)$ en $\alpha'\underline{x} \sim N(\alpha'\mu,\alpha'\Sigma^2\alpha)$ voor alle $\alpha \neq 0$. Dus

$$\varphi_{\alpha'\underline{x}}(t) = \exp(it\alpha'\mu - \tfrac{1}{2}\alpha'\Sigma^2\alpha t^2) \;;$$

stel $t = 1$ dan is

$$\varphi_{\alpha'\underline{x}}(t) = \varphi_{\underline{x}}(\alpha t) = \varphi_{\underline{x}}(\alpha) = \exp(i\alpha'\mu - \tfrac{1}{2}\alpha'\Sigma^2\alpha)$$

dus $\underline{x} \sim N_p(\mu,\Sigma^2)$.

Uit theorema 3.3 volgt ook direkt:

3.8. Theorema

Is $\underline{x} \sim N_p(\mu,\Sigma^2)$ dan is de *marginale verdeling* van elke deelverzameling van $\underline{x}$ weer normaal met corresponderende gemiddelde en variantie-covariantiematrix.

Bewijs: Laat $\underline{x}' = (\underline{x}^{(1)'}, \underline{x}^{(2)'}) = (\underline{x}_1, \ldots, \underline{x}_{p_1}, \underline{x}_{p_1+1}, \ldots, \underline{x}_p)'$;

$\mu' = (\mu^{(1)'}, \mu^{(2)'})$ en $\Sigma^2 = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$. Hierin is dus $\Sigma_{11} = \mathrm{VAR}\ \underline{x}^{(1)}$,

$\Sigma_{22} = \mathrm{VAR}\ \underline{x}^{(2)}$ (beide regulier ondersteld) en $\Sigma_{12} = \mathrm{COV}(\underline{x}^{(1)}, \underline{x}^{(2)}) = \Sigma_{21}'$.

Neem nu $A_{p_1 \times p} = (I_{p_1 \times p_1}, 0)$ dan volgt $\underline{x}^{(1)} = A\underline{x} \sim N_{p_1}(\mu^{(1)}, \Sigma_{11})$.

(Merk op dat in $\Sigma_{11}$ het kwadraat ontbreekt; eventueel te schrijven als $\Sigma_1^2$.)

3.9. Theorema

Is $\underline{x} \sim N_p(\mu, \Sigma^2)$, dan geldt: $\underline{x}^{(1)}, \underline{x}^{(2)}$ *ongecorreleerd* $\Longleftrightarrow$ $\underline{x}^{(1)}, \underline{x}^{(2)}$ *stochastisch onafhankelijk*.

Bewijs: Dit volgt direkt uit de karakteristieke funktie, nl. $\Sigma_{12} = \Sigma_{21}' = 0$ geeft:

$$\varphi_{\underline{x}}(t) = \exp[it^{(1)'}\mu^{(1)} + it^{(2)'}\mu^{(2)} - \tfrac{1}{2}t^{(1)'}\Sigma_{11}t^{(1)} - \tfrac{1}{2}t^{(2)'}\Sigma_{22}t^{(2)}] =$$

$$= \varphi_{\underline{x}^{(1)}}(t^{(1)})\varphi_{\underline{x}^{(2)}}(t^{(2)}),$$ dus $\underline{x}^{(1)}$ en $\underline{x}^{(2)}$ stochastisch onafhankelijk (zie eind 2.44).

Hierin is $t' = (t^{(1)'}, t^{(2)'}) = (t_1, \ldots, t_{p_1}, t_{p_1+1}, \ldots, t_p)'$.

Pas op: ongecorreleerde, simultaannormaal verdeelde variabelen zijn onafhankelijk. Zijn ze echter niet simultaannormaal verdeeld dan is dat niet zonder meer juist (zie appendix A).

3.10. Theorema

Is $\underline{x} \sim N_p(\mu, \Sigma^2)$ dan is de *voorwaardelijke verdeling* van $\underline{x}^{(1)}$ bij gegeven $x^{(2)}$

$$f(x^{(1)} \mid x^{(2)}) = N_{p_1}[\mu^{(1)} + \mathcal{B}(x^{(2)} - \mu^{(2)}), \Sigma_{11\cdot 2}]$$

met $\mathcal{B} := \Sigma_{12}\Sigma_{22}^{-1}$ en $\Sigma_{11\cdot 2} := \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ (zie 1.35).

$\mathcal{B}$ heet de *matrix der regressiecoëfficiënten* van $\underline{x}^{(1)}$ op $x^{(2)}$ (beter $\mathcal{B}_{12}$).

Bewijs: Voer een reguliere transformatie uit $\underline{y} = A\underline{x}$ zodat $\underline{y}^{(1)}$ en $\underline{y}^{(2)}$ ongecorreleerd zijn

$$\begin{aligned}\underline{y}^{(1)} &= \underline{x}^{(1)} - \mathcal{B}\underline{x}^{(2)} \\ \underline{y}^{(2)} &= \underline{x}^{(2)}\end{aligned} \qquad \text{dus} \quad A = \begin{pmatrix} I_{p_1\times p_1} & -\mathcal{B} \\ 0 & I \end{pmatrix} \; ; \quad |A| = 1 \; .$$

Dus

$$\mathrm{COV}(\underline{x}^{(1)} - \mathcal{B}\underline{x}^{(2)}, \underline{x}^{(2)}) = \Sigma_{12} - \mathcal{B}\,\Sigma_{22} = 0 \quad \text{oftewel} \quad \mathcal{B} = \Sigma_{12}\Sigma_{22}^{-1} \; .$$

$$\mathrm{VAR}\;\underline{y}^{(1)} = \Sigma_{11} + \mathcal{B}\Sigma_{22}\mathcal{B}' - \mathcal{B}\Sigma_{21} - \Sigma_{12}\mathcal{B}' = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = \Sigma_{11\cdot 2}$$

en

$$\mathcal{E}\underline{y}^{(1)} = \mu^{(1)} - \mathcal{B}\mu^{(2)} \; .$$

Nu is $\underline{x}$, dus volgens 3.3 ook $\underline{y}$, eveneens $\underline{y}^{(1)}$ volgens 3.8 normaal verdeeld. $\underline{y}^{(1)}$ en $\underline{y}^{(2)}$ zijn ongecorreleerd, dus volgens 3.9 ook onafhankelijk, dus $f(y^{(1)} \mid y^{(2)}) = f(y^{(1)})$.

Daar $f(x^{(1)} \mid x^{(2)}) = f(y^{(1)} + \mathcal{B}x^{(2)} \mid x^{(2)})$ is f een normale verdeling met $\mathcal{E}[\underline{x}^{(1)} \mid x^{(2)}] =$
$= \mathcal{E}\underline{y}^{(1)} + \mathcal{B}x^{(2)} = \mu^{(1)} - \mathcal{B}\mu^{(2)} + \mathcal{B}x^{(2)}$ en $\mathrm{VAR}[\underline{x}^{(1)} \mid x^{(2)}] = \mathrm{VAR}\;\underline{y}^{(1)} =$
$= \Sigma_{11\cdot 2}$. Dus

$$f(x^{(1)} \mid x^{(2)}) = N_{p_1}[\mu^{(1)} + \mathcal{B}(x^{(2)} - \mu^{(2)}), \Sigma_{11\cdot 2}] \; .$$

Enkele opmerkingen:

a) We zien inderdaad, indien $\Sigma_{12} = 0$, dus $\mathcal{B} = 0$ en $\Sigma_{11\cdot 2} = \Sigma_{11}$, dat $f(x^{(1)} \mid x^{(2)}) = N_{p_1}(\mu^{(1)}, \Sigma_{11}) = f(x^{(1)})$ ofwel ongecorreleerd $\equiv$ onafhankelijk.

b) Het geval p = 2 geeft:

3.11. $$f(x_1 \mid x_2) = N\left[\mu_1 + \frac{\sigma_1}{\sigma_2}\rho(x_2 - \mu_2), \sigma_1^2(1 - \rho^2)\right] \; .$$

c) $\mathrm{VAR}[\underline{x}^{(1)} \mid x^{(2)}] = \Sigma_{11\cdot 2}$ is *onafhankelijk* van het niveau van $x^{(2)}$. Dit is essentieel en geldt niet zonder meer voor andere verdelingen.

3.12. d) Is $\mu = 0$, dan is $\mathcal{E}[\underline{x}^{(1)}|x^{(2)}] = \mathcal{B}x^{(2)}$ en $\mathcal{E}[\underline{x}_i|x^{(2)}] = \beta_{i*}' \; x^{(2)}$.

Stellen we nu $\underline{x}_i = \underline{y}$, $x^{(2)} = (x_0,\ldots,x_p)' = x$, $\beta_{i*}' = \beta' = (\beta_0,\ldots,\beta_p)$, dan wordt dit juist

$$\mathcal{E}[\underline{y}|x] = \beta' x = \beta_0 x_0 + \ldots + \beta_p x_p$$

of ook geschreven

$$\underline{y} = \beta' x + \underline{e} \quad \text{met } \mathcal{E}\underline{e} = 0 \quad \text{en} \quad \text{var } \underline{e} = \sigma^2 \text{ (onafhank. x).}$$

Dit is een rechtvaardiging voor het feit dat dit *lineaire model* in de multiple regressie analyse vaak wordt toegepast. In de regressie analyse schat men o.a. de *regressiecoëfficiënten* $\beta_{i*}$, dat is dus één rij uit de *matrix der regressiecoëfficiënten* $\mathcal{B}$ (zie § 5 multivariate analyse).

e) $\Sigma_{11.2} > 0$ nl. het is een variantie-covariantie matrix, dus zeker $\Sigma_{11.2} \geq 0$.

Echter $\Sigma_{11.2}$ is regulier als $\Sigma^2$ regulier is, nl

$$0 < |\Sigma^2| = |\Sigma_{22}||\Sigma_{11.2}| \quad \text{dus } |\Sigma_{11.2}| \neq 0$$

f) Er geldt $\Sigma_{11.2} \leq \Sigma_{11}$.

## § 4. De verdeling van kwadratische vormen

### a. De niet-centrale $\chi^2$-, F- en t-verdeling

4.1. Definitie: $\chi^2_\nu(\mu^2) := (\mu + \underline{\chi}_\nu)^2$ is een niet-centrale chikwadraatvariable met $\nu$ vrijheidsgraden en niet-centraliteitsparameter $\delta := \mu^2$.

$\underline{\chi}_\nu \sim N_\nu(0,I)$.

Voor $\delta$ (dus $\mu$) = 0 is dit de centrale chikwadraat.

Uit 2.36 volgt direct:

4.2. $\mathcal{E}\underline{\chi}^2_\nu(\delta) = \nu + \delta$ en uit 4.7ii) var $\underline{\chi}^2_\nu(\delta) = 2\nu + 4\delta$.

4.3. De momentgenererende functie van $\underline{v} := \chi^2_\nu(\delta)$ luidt:

$$m_{\underline{v}}(t) = (1 - 2t)^{-\nu/2}\exp[\delta t/(1 - 2t)] .$$

Bewijs: Hiertoe leiden we eerst algemener af:

Zij $\underline{y} \sim N_n(0,I)$, C symmetrisch en $\underline{z} := \underline{y}'C\underline{y} + m'\underline{y} + c$

De momentgenererende functie van $\underline{z}$ is

4.4. $$m_{\underline{z}}(t) = |I - 2tC|^{-\frac{1}{2}}\exp[\tfrac{1}{2}t^2m'(I - 2tC)^{-1}m + c]$$

$$m_{\underline{z}}(t) = \mathcal{E}e^{t\underline{z}} = (2\pi)^{-\frac{1}{2}n}\int_{-\infty}^{\infty}\!\!\dots\!\int_{-\infty}^{\infty}\exp(ty'Cy + tm'y + tc - \tfrac{1}{2}y'y)dy_1\dots dy_n$$

Dit is volgens vraagstuk 3.13, met A = 0, a = o, $\alpha$ = 1, $-b = tm$, $-\beta = tc$ en $B = \frac{1}{2}I - tC$

$$m_{\underline{z}}(t) = 2^{-\frac{1}{2}n}|\tfrac{1}{2}I - tC|^{-\frac{1}{2}}\exp[\tfrac{1}{4}t^2m'(\tfrac{1}{2}I - tC)^{-1}m + tc] =$$

$$= |I - 2tC|^{-\frac{1}{2}}\exp[\tfrac{1}{2}t^2m'(I - 2tC)^{-1}m + tc]. \qquad \square$$

Opmerking: $I - 2tC > 0$ voor voldoend kleine t, zie vraagstuk 1.41.

Stel nu $\underline{x} = \mu + \underline{y}$, dus $\underline{v} = \underline{x}'\underline{x} \sim \chi^2_\nu(\delta)$

$\underline{x}'\underline{x} = (\mu + \underline{y})'(\mu + \underline{y}) = \underline{y}'\underline{y} + 2\mu'\underline{y} + \mu'\mu$. Neem C = I, $m = 2\mu$, $c = \mu'\mu$, dan

$m_{\underline{v}}(t) = (1 - 2t)^{-\frac{1}{2}\nu}\exp[2t^2\mu'\mu(1 - 2t)^{-1} + t\mu'\mu]$ en $\mu'\mu = \delta$ geeft

$$m_{\underline{v}}(t) = (1 - 2t)^{-\frac{1}{2}\nu}\exp[\delta t/(1 - 2t)] \qquad \square$$

4.5. Stelling: Als $\underline{v}_1 \simeq \chi^2_{\nu_1}(\delta_1)$ onafhankelijk $\underline{v}_2 \simeq \chi^2_{\nu_2}(\delta_2)$, dan is

$$\underline{v}_1 + \underline{v}_2 \simeq \chi^2_{\nu_1+\nu_2}(\delta_1 + \delta_2) .$$

Bewijs:

$m_{\underline{v}_1+\underline{v}_2}(t) = m_{\underline{v}_1}(t)m_{\underline{v}_2}(t) =$

$$= (1-2t)^{-\nu_1/2}\exp[\delta_1 t/(1-2t)](1-2t)^{-\nu_2/2}\exp[\delta_2 t/(1-2t)] =$$

$$= (1-2t)^{-(\nu_1+\nu_2)/2}\exp[(\delta_1+\delta_2)t/(1-2t)] \;,$$

(analoog 2.44).

4.6. Definitie: $\underline{t}_\nu(\mu) := \dfrac{\mu+\underline{\chi}}{\sqrt{\underline{\chi}_\nu^2/\nu}}$, $\underline{\chi}$ onafhankelijk $\underline{\chi}_\nu$, is een niet-centrale student-variabele met $\nu$ vrijheidsgraden en niet-centraliteitsparameter $\mu$.

4.7. Definitie: $\underline{F}_{\nu_2}^{\nu_1}(\delta) := \dfrac{\underline{\chi}_{\nu_1}^2(\delta)/\nu_1}{\underline{\chi}_{\nu_2}^2/\nu_2}$, $\underline{\chi}_{\nu_1}$ onafhankelijk $\underline{\chi}_{\nu_2}$, is een niet-centrale F-variabele met $\nu_1$ resp. $\nu_2$ vrijheidsgraden en niet-centraliteitsparameter $\delta$. Voor de kansdichtheid, variantie e.d. wordt verwezen naar Searle.

b. Kwadratische vormen

4.8. $\underline{x} \sim N_p(\mu,\Sigma^2) \Rightarrow \mathrm{var}(\underline{x}'A\underline{x}) = 2\,\mathrm{sp}[(A\Sigma^2)^2] + 4\mu'A\Sigma^2A\mu.$

Bewijs: Stel $\underline{x} = \mu + \Sigma'H\underline{z}$ met H zodat $H'\Sigma A\Sigma'H = \Lambda = \mathrm{diag}(\lambda_1,\dots,\lambda_p)$. Dus
$\underline{z} \sim N_p(0,I)$ en var $\underline{z}_i^2 = 2$.
$\underline{x}'A\underline{x} = (\mu' + \underline{z}'H'\Sigma)A(\mu + \Sigma'H\underline{z}) = \mu'A\mu + 2\mu'A\Sigma'H\underline{z} + \underline{z}'\Lambda\underline{z}.$
$\mathrm{var}(\underline{x}'A\underline{x}) = 4\mu'A\Sigma^2A\mu + \mathrm{var}(\underline{z}'\Lambda\underline{z}) + 4\mu'A\Sigma'H\;\mathrm{cov}(\underline{z},\underline{z}'\Lambda\underline{z}).$
Nu is
$\mathrm{var}(\underline{z}'\Lambda\underline{z}) = \mathrm{var}\;\Sigma\,\lambda_i\underline{z}_i^2 = 2\;\Sigma\,\lambda_i^2 = 2\,\mathrm{sp}\,[\Lambda^2] = 2\mathrm{sp}\,[(\Sigma A\Sigma')^2] = 2\mathrm{sp}[(A\Sigma^2)^2].$
$\mathrm{cov}(\underline{z},\underline{z}'\Lambda\underline{z}) = \mathcal{E}\,[\underline{z}(\underline{z}'\Lambda\underline{z} - \mathrm{sp}\,\Lambda)] = \mathcal{E}(\underline{z}\underline{z}'\Lambda\underline{z}) = \mathcal{E}(\underline{z}\Sigma\lambda_i\underline{z}_i^2) =$
$= \mathcal{E}(\underline{z}_1\Sigma\lambda_i\underline{z}_i^2,\dots,\underline{z}_p\Sigma\lambda_i\underline{z}_i^2)' = 0$ daar $\mathcal{E}(\underline{z}_i\underline{z}_j^2) = 0$ en $\mathcal{E}\underline{z}_i^3 = 0.$

Gevolg:

i) $p = 1$, $A = 1$, $\Sigma^{\underline{2}} = \sigma^2$ geeft $\underline{x} \sim N(\mu,\sigma^2)$ en var $\underline{x}^2 = 2\sigma^4 + 4\sigma^2\mu^2$.

ii) $p = \nu$, $A = \Sigma^{\underline{2}} = I$ geeft $\underline{x} = \mu + \underline{\chi}_\nu$ en var $\underline{x}^{\underline{2}} = 2\nu + 4\mu^2$.

4.9. $\mathrm{cov}(B\underline{x},\underline{x}'A\underline{x}) = 2B\Sigma^{\underline{2}}A\mu$.

Bewijs:

$$B\ \mathrm{cov}(\underline{x},\underline{x}'A\underline{x}) = B\ \mathrm{cov}(\mu + \Sigma'H\underline{z},\ \mu'A\mu + 2\mu'A\Sigma'H\underline{z} + \underline{z}'\Lambda\underline{z}) =$$

$$= B\ \mathrm{cov}(\Sigma'H\underline{z}, 2\mu'A\Sigma'H\underline{z}) = 2B\Sigma'HH'\Sigma A\mu = 2B\Sigma^{\underline{2}}A\mu\ .$$

4.10. Hulpstelling: $C$ symmetrisch; $\underline{y} \sim N_n(0,I)$

$\underline{z} := \underline{y}'C\underline{y} + m'\underline{y} + c \sim \underline{\chi}^2_\nu(\delta)$ met $\delta = c$, $\nu = r(C)$

d.e.s.d. als i) $C^2 = C$

ii) $Cm = m$

iii) $c = \frac{1}{4}m'm$

Bewijs:

Dus de m.g. functies zijn gelijk:

$$(1-2t)^{-\frac{1}{2}\nu}\exp[\delta t(1-2t)^{-1}] = |I - 2tC|^{-\frac{1}{2}}\exp[\tfrac{1}{2}t^2m'(I-2tC)^{-1}m + tc]$$

Ofwel

$1^e$ $\quad (1-2t)^\nu = |I - 2tC|$

voor alle t

$2^e$ $\quad \delta t = tc(1-2t) + \frac{1}{2}t^2m'(I - 2tC)^{-1}m(1-2t)$

$1^e$ Stel $C = H\ \Lambda\ H'$ en $2t = \tau$; $|I - 2tC| = |I - \tau\ H\ \Lambda\ H'| = \prod_1^n(1-\tau\lambda_i) = (1-\tau)^\nu$

Beide polynomen in $\tau$ zijn gelijk, dus $C$ heeft $\nu$ eigenw. 1 en $n - \nu$ eigenw. 0 dus volgens (1.93) is $C$ idempotent en $r(C) = \nu$.

$2^e$ $(I - 2tC)^{-1} = \sum_0^\infty (2tC)^k = I + 2tC + \ldots$ zie (1.77).

$$\delta t = tc - 2t^2c + \tfrac{1}{2}t^2m'(I + 2tC + \ldots)m(1-2t)$$

$$0 = t(c-\delta) - t^2(2c - \tfrac{1}{2}m'm) + t^3(m'Cm - m'm) + \ldots$$

Dus $\delta = c = \frac{1}{4}m'm$ en $m'Cm = m'm$ en daar $C^2 = C$, $Cm = m$. □

4.11. Stelling: $\underline{x} \sim N_p(\mu,\Sigma^2)$; $\Sigma^2 =: V \geq 0$ dus singulier toegelaten
A niet noodzakelijk symmetrisch, VAV symmetrisch voldoende

$\underline{x}'A\underline{x} \sim \underline{\chi}^2_\nu(\mu'A\mu)$ met $\nu = sp(AV)$ d.e.s.d. als i) $VAVAV = VAV$
ii) $VAVA\mu = VA\mu$
iii) $\mu'A\mu = \mu'AVA\mu$ $\Big\}$ A symmetrisch ondersteld

Opmerking: is $\Sigma^2 > 0$ dan noodzakelijk en voldoende AV idempotent en $\nu = r(A)$

Bewijs: Stel $\underline{x} = \mu + \Sigma\underline{y}$; $\underline{x}'A\underline{x} = \underline{y}'\Sigma A\Sigma\underline{y} + 2\mu'A\Sigma\underline{y} + \mu'A\mu$.
Dus i) $\Sigma A\Sigma\Sigma A\Sigma = \Sigma A\Sigma$ of met (1.60) $VAVAV = VAV$
ii) $\Sigma A\Sigma\Sigma A\mu = \Sigma A\mu$ oftewel $VAVA\mu = VA\mu$
iii) $\mu'A\mu = \frac{1}{4}\cdot 4\mu'A\Sigma\Sigma A\mu = \mu'AVA\mu$.

Opmerking: In het bewijs is $A' = A$ ondersteld; is dit niet zo dan worden ii) en iii) ingewikkelder.

Belangrijke toepassing:

4.12. Is P een orthogonale projector met $r(P) = r$ en zij $\underline{x} \sim N_p(\mu,\sigma^2 I)$ dan is
$\underline{x}'P\underline{x}/\sigma^2 = \|P\underline{x}\|^2/\sigma^2 \sim \underline{\chi}^2_r(\delta)$ met $\delta = \|P\mu\|^2/\sigma^2$.

4.13. Stelling: $\underline{x} \sim N_p(\mu,I)$; $\underline{x}'A\underline{x}$ onafhankelijk $\underline{x}'B\underline{x} \Leftrightarrow BA = 0$ (A, B symmetrisch).

Bewijs:

$\Rightarrow$ : Dus $m_{\underline{x}'A\underline{x}+\underline{x}'B\underline{x}}(t) = m_{\underline{x}'(A+B)\underline{x}}(t) = m_{\underline{x}'A\underline{x}}(t)m_{\underline{x}'B\underline{x}}(t)$, dus

$$|I - 2t(A+B)| = |I - 2tA||I - 2tB|$$

dus volgens 1.82 moet $AB = 0$.

$\Leftarrow$ : Volgens 1.83 is er een H zodat $H'AH = \Lambda_1$ en $H'BH = \Lambda_2$.
$\Lambda_1\Lambda_2 = 0$ d.w.z. $\lambda_{i1}\lambda_{i2} = 0 \forall_i$ dus minstens één van beide 0.
Stel $\underline{x} = H\underline{z}$ dus $\underline{z} \sim N_p(\mu_z, I)$.
$\underline{x}'A\underline{x} = \underline{z}'\Lambda_1\underline{z} = \Sigma\lambda_{i1}\underline{z}_i^2$ en $\underline{x}'B\underline{x} = \underline{z}\Lambda_2\underline{z} = \Sigma\lambda_{i2}\underline{z}_i^2$ dus onafhankelijk.

4.14. Stelling: $\underline{x} \sim N_p(\mu, I)$, A symmetrisch.

$B\underline{x}$ onafhankelijk $\underline{x}'A\underline{x} \Leftrightarrow BA = 0$

Bewijs:

$\Leftarrow$ : Stel $\underline{x} = H\underline{z}$ dus $\underline{z} \sim N_p(\mu_z, I)$. Dan is $B\underline{x} = BH\underline{z}$

$\underline{x}'A\underline{x} = \underline{z}'H'AH\underline{z} = \underline{z}'\Lambda\underline{z}$. Nu is $BA = 0 = BHH'A$, dus ook $BHH'AH = BH\Lambda =: C\Lambda = 0$.

Dus

$$C\Lambda = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \begin{pmatrix} \Lambda_1 & 0 \\ 0 & 0 \end{pmatrix} = 0$$

waaruit volgt $C_{11} = 0$, $C_{21} = 0$ oftewel

$$C = \begin{pmatrix} 0 & C_{12} \\ 0 & C_{22} \end{pmatrix} =: (0 \quad C_2) .$$

Stel $\underline{z} = (\underline{z}_1', \underline{z}_2')'$.

Nu is $B\underline{x} = C\underline{z} = C_2\underline{z}_2$ en $\underline{x}'A\underline{x} = \underline{z}_1\Lambda_1\underline{z}_1$.

Daar $\underline{z}_1$ en $\underline{z}_2$ onafhankelijk zijn, dus ook $B\underline{x}$ en $\underline{x}'A\underline{x}$.

$\Rightarrow$ : Dus ook $\underline{x}'B'B\underline{x}$ o.o. $\underline{x}'A\underline{x}$ dus volgens 4.13 geldt $B'BA = 0$ en met (1.60) $BA = 0$.

Gevolg:

4.13*. Stelling: $\underline{x} \sim N_p(\mu, \Sigma^2)$ met $\Sigma^2 =: V \geq 0$.

$$\underline{x}'A\underline{x} \text{ o.o. } \underline{x}'B\underline{x} \Leftrightarrow \begin{cases} VAVBV = 0 \\ VAVB\mu = VBVA\mu = 0 \\ \mu'AVB\mu = 0 \end{cases}$$

Is $\Sigma^2$ regulier, dan noodzakelijk en voldoende $A\Sigma^2B = 0$.

Bewijs:

Stel $\underline{x} = \mu + \Sigma\underline{y}$ met $\underline{y} \sim N_p(0, I)$

$\underline{x}'A\underline{x} = \underline{y}'\Sigma A\Sigma\underline{y} + 2\mu'A\Sigma\underline{y} + \mu'A\mu$ analoog $\underline{x}'B\underline{x}$.

$1^e$) $\underline{y}'\Sigma A\Sigma\underline{y}$ o.o. $\underline{y}'\Sigma B\Sigma\underline{y}$ volgens 4.13 d.e.s.d. als $\Sigma A\Sigma\Sigma B\Sigma = 0$ oftewel met (1.96) $VAVBV = 0$

$2^e$) $\mu'A\Sigma\underline{y}$ o.o. $\underline{y}'\Sigma B\Sigma\underline{y}$ volgens 4.14 d.e.s.d. als $\mu'A\Sigma\Sigma B\Sigma = 0$ oftewel $\mu'AVBV = 0$, dus $VBVA\mu = 0$. Analoog $VAVB\mu = 0$.

$3^e$) $\mu'A\Sigma\underline{y}$ o.o. $\mu'B\Sigma\underline{y}$; $\mathrm{cov}(\mu'A\Sigma\underline{y}, \mu'B\Sigma\underline{y}) = \mu'A\Sigma\Sigma B\mu = \mu'AVB\mu = 0$.

Opmerking: Voldoende is weer dat VAV en VBV symmetrisch zijn.

Gevolg:

4.14.* Stelling: $\underline{x} \sim N_p(\mu,\Sigma^2)$ met $\Sigma^2 =: V \geq 0$.

$$\underline{x}'A\underline{x} \text{ o.o. } B\underline{x} \Leftrightarrow \begin{cases} BVAV = 0 \\ BVA\mu = 0. \end{cases}$$

Is $\Sigma^2$ regulier, dan noodzakelijk en voldoende $B\Sigma^2A = 0$

Bewijs:

Stel $\underline{x} = \mu + \Sigma\underline{y}$ met $\underline{y} \sim N_p(o,I)$.

$$\underline{x}'A\underline{x} = \underline{y}'\Sigma A\Sigma\underline{y} + 2\mu'A\Sigma\underline{y} + \mu'A\mu \quad \text{en } B\underline{x} = B\mu + B\Sigma\underline{y}.$$

$1^e$) $B\Sigma\underline{y}$ o.o. $\underline{y}'\Sigma A\Sigma\underline{y}$ volgens 4.14 $B\Sigma\Sigma A\Sigma = 0$ oftewel $BVAV = 0$.

$2^e$) $B\Sigma\underline{y}$ o.o. $\mu'A\Sigma\underline{y}$; $\mathrm{cov}(B\Sigma\underline{y}, \mu'A\Sigma\underline{y}) = B\Sigma\Sigma A\mu = BVA\mu = 0$.

Voldoende is weer dat VAV symmetrisch is.

4.15. Belangrijke toepassing:

Is $\underline{x} \sim N_p(\mu,\sigma^2 I)$. Zijn $P_1$ en $P_2$ orthogonale projecties op onderling loodrechte deelruimten, dan is $P_1P_2 = 0$ en zijn $P_1\underline{x}$ en $\underline{x}'P_2\underline{x}$ onafhankelijk.

## § 5. Regressie, het model van volle rang

### a. Het model

Regressie-analyse is ontworpen voor situaties waar een variabele y (afhankelijke variabele) samenhangt met één of meer andere variabelen $x_i$ (onafhankelijke variabelen) die gewoonlijk alle gemeten worden aan hetzelfde object. Doel is met deze metingen die samenhang te vinden of te controleren.

Voorbeeld: Uit informatie over het inkomen y en het aantal jaren scholing x proberen we te achterhalen in welke mate y gerelateerd is aan x. Omdat het inkomen van een individu ongetwijfeld nog door andere (onbekende) factoren dan de scholing zal worden beïnvloed en er toevallige storingen bij de registratie van y kunnen optreden, is het streven gericht op het vinden van een relatie tussen het gemiddelde, verwachte inkomen van een individu met aantal jaren scholing x (bekend) en x zelf. Het inkomen in een concreet geval zal van dit gemiddelde afwijken. Ongeacht de verschillende oorzaken van deze afwijking vatten we deze samen in een stochastische storingsterm, de toevallige fout $\underline{e} := \underline{y} - \mathcal{E}\underline{y}$. Voor de relatie tussen $\mathcal{E}\underline{y}$ en x is in principe alles mogelijk.

Algemeen: $\mathcal{E}\underline{y} = f(x)$ oftewel $\underline{y} = f(x) + \underline{e}$ heet een regressiemodel. We beperken ons tot lineaire regressiemodellen omdat hiervoor de theorie volledig is en het practische nut bewezen.

De lineariteit slaat op de parameters bv.: $\mathcal{E}\underline{y} = \alpha + \beta x$, $\mathcal{E}\underline{y} = \alpha + \beta e^x$, $\mathcal{E}\underline{y} = \alpha + \beta x + \gamma x^2$, doch niet $\mathcal{E}\underline{y} = \alpha + x^\beta$.

De verklarende variabele x wordt als niet stochastisch beschouwd, als "instelbaar" op iedere vooraf te bepalen waarde.

In het vervolg nemen we als model $\mathcal{E}\underline{y} = \alpha + \beta x$ ($\mathcal{E}\underline{y} = \alpha + \beta e^x = \alpha + \beta z$ e.d. zijn hiertoe te herleiden!).

### b. Waarnemingen

Er worden aselect n onafhankelijke individuen gekozen, elk met een verschillend aantal jaren scholing $x_i$. Hun inkomen wordt eveneens geregistreerd.

De afwijking van $y_i$ van het bij deze $x_i$ geldende gemiddelde $\mathcal{E}\underline{y}_i = \alpha + \beta x_i$ is $e_i := y_i - \mathcal{E}y_i$.

Het model luidt nu: $\underline{y}_i = \alpha + \beta x_i + \underline{e}_i$ met $\mathcal{E}\underline{e}_i = 0$ var $\underline{e}_i = \sigma^2$ en $\mathrm{cov}(\underline{e}_i, \underline{e}_j) = 0$ $(i \neq j)$, $i = 1,\ldots,n$.

c. Schatters voor de onbekende parameters $\alpha$ en $\beta$

De hier gebruikte methode der kleinste kwadraten houdt in dat $\alpha$ en $\beta$ zo geschat dienen te worden, dat de som van de kwadraten van de afwijkingen van de waargenomen $y_i$'s van hun gemiddelde $\mathcal{E}\underline{y}_i$ zo klein mogelijk is:

$$f := e'e = \sum_1^n e_i^2 = \Sigma(y_i - \mathcal{E}\underline{y}_i)^2 = \Sigma(y_i - \alpha - \beta x_i)^2 \ .$$

Dit minimum, bereikt voor a en b, vinden we als volgt:

$$\frac{\delta f}{\delta \alpha} = -2(\Sigma y_i - n\alpha - \beta\Sigma x_i) = 0 \text{ geeft } na + b\Sigma x_i = \Sigma y_i$$

$$\frac{\delta f}{\delta \beta} = -2(\Sigma x_i y_i - \alpha\Sigma x_i - \beta\Sigma x_i^2) = 0 \text{ geeft } a\Sigma x_i + b\Sigma x_i^2 = \Sigma x_i y_i \ .$$

Hieruit a en b oplossen geeft:

5.1. $$a = \bar{y}_. - b\bar{x}_., \quad b = \frac{\Sigma(x_i - \bar{x}_.)(y_i - \bar{y}_.)}{\Sigma(x_i - \bar{x}_.)^2} = \frac{\Sigma x_i y_i - n\bar{x}_.\bar{y}_.}{\Sigma x_i^2 - n\bar{x}_.^2}$$

waarin $\bar{x}_. := \sum_1^n x_i/n$ en analoog $\bar{y}_. := \sum_1^n y_i/n$.

Dat deze waarden inderdaad een minimum opleveren, zien we als volgt:
$\partial^2 f/\partial\alpha^2 = 2n$; $\partial^2 f/\partial\beta^2 = 2\Sigma x_i^2$ en $\partial^2 f/\partial\alpha\partial\beta = 2\Sigma x_i$, zodat

$$\det H = \begin{vmatrix} 2n & 2\Sigma x_i \\ 2\Sigma x_i & 2\Sigma x_i^2 \end{vmatrix} = 4[n\Sigma x_i^2 - (\Sigma x_i)^2] = 4\Sigma(x_i - \bar{x}_.)^2 > 0 \ ,$$

dus we hebben met een extremum te maken en omdat de elementen op de hoofddiagonaal van H positief zijn is het een minimum.

d. Meerdere onafhankelijke (verklarende) variabelen

Uitbreiding van het voorbeeld:
Naast de variabele "aantal jaren scholing" wordt de variabele "leeftijd" als invloedsfactor van het inkomen in het model betrokken.

We nemen dan aan $\mathcal{E}\underline{y} = \alpha + \beta_1 x_1 + \beta_2 x_2$ waarin $x_1$ het aantal jaren scholing en $x_2$ de leeftijd is. Van n individuen met een verschillende $x_1,x_2$-combinatie wordt het inkomen genoteerd. Er geldt:

$$\underline{y}_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \underline{e}_i \ , \qquad i = 1,\dots,n \ .$$

Noteer $\alpha = \beta_0 x_{i0}$ dus $\beta_0 = \alpha$ en $x_{i0} = 1$ voor $i = 1,\dots,n$.
Nu algemeen:
Een variabele y wordt gerelateerd aan k onafhankelijke verklarende variabelen $x_1,\dots,x_k$.
Het model wordt nu in matrixvorm:

$$\underline{y} = X\beta + \underline{e} \quad \text{met} \quad \mathcal{E}\underline{e} = 0 \quad \text{en} \quad \text{var}\ \underline{e} = \sigma^2 I_n \ .$$

Hierin is dus:

$\underline{y} = (y_1,\dots,y_n)'$ de vector der waarnemingen, response vector
$\beta = (\beta_0,\dots,\beta_k)'$ de vector der parameters
$\underline{e} = (\underline{e}_1,\dots,\underline{e}_n)'$ de storingsvector (random disturbance)
$X = (u\ x_{*1} \dots x_{*k})$ de matrix der instellingen, design matrix.

Opmerking:
i) $n > k$ ondersteld omdat er anders geen schattingsprobleem voor de parameters $\beta$ bestaat. Dan kunnen ze zo bepaald worden dat $y = X\beta$.
ii) De aanname var $\underline{e} = \sigma^2 I$, d.w.z. de fouten zijn ongecorreleerd en hebben dezelfde variantie, is een essentiële beperking van het model!

Zij $\underline{b}$ de kleinste kwadratenschatter voor $\beta$, eveneens $\hat{\underline{y}} := X\underline{b}$ voor $\mathcal{E}\underline{y} = X\beta$. Deze vinden we weer door $e'e$ te minimaliseren:

5.2. $$\min(e'e) = \|y - \hat{y}\|^2 = \|y - Xb\|^2 = \min_{\beta}\|y - X\beta\|^2 \ .$$

Nu is

$$\|y - X\beta\|^2 = (y - X\beta)'(y - X\beta) = y'y - y'X\beta - \beta'X'y + \beta'X'X\beta$$
$$= y'y - 2\beta'X'y + \beta'X'X\beta \ .$$

$$\frac{\partial(e'e)}{\partial\beta'} = 0 = -2X'y + 2X'X\beta$$

(zie B8) oftewel

5.3. $$X'X\underline{b} = X'\underline{y} \ , \ \text{de normaalvergelijkingen} \ .$$

Als $X'X$ regulier is, d.w.z. $r(X) = k + 1$, is dit stelsel éénduidig oplosbaar. Door de keuze van de instellingen $x_{ij}$ kan men dit garanderen, d.w.z. dat $X$ van volle kolomrang is (vandaar de naam "model van volle rang").

5.4. De schatter voor $\beta$ is dan $\underline{b} = (X'X)^{-1}X'\underline{y}$.

5.5. En voor $X\beta = \mathcal{E}\underline{y}$: $\qquad \hat{\underline{y}} = Xb = X(X'X)^{-1}X'\underline{y} =: P_X\underline{y}$ (zie 1.91).

<u>Opmerking</u>: $d(y,\langle X\rangle) = \|y - \hat{y}\|$ is dus de (minimale) afstand van $y$ tot $\langle X\rangle$
$\hat{y} = Xb \in \langle X\rangle$ (zie 1.9).

e. <u>Correcties t.o.v. het gemiddelde</u>

Ondermeer op numerieke gronden (de elementen van $X'X$ kunnen erg groot zijn) worden in de practijk de waarnemingen $\underline{y}$ en de x-variabelen gecorrigeerd door het gemiddelde ervan af te trekken.
We partitioneren en voeren de volgende notaties in:

5.6. $\beta = (\beta_0,\beta_1,\dots,\beta_k)' =: (\beta_0,\beta_m')'$; analoog $b_m = (b_1,\dots,b_k)'$.

$X = (u\ X_1)$ d.w.z. $X_1 := \{x_{ij}\}$ met $i = 1,\dots,n$ en $j = 1,\dots,k$.

$u := (1,\dots,1)'$ een vector met $n$ componenten 1;

$\bar{x} := (\bar{x}_{\cdot 1},\dots,\bar{x}_{\cdot k})'$ waarin $\bar{x}_{\cdot j} := \sum_i x_{ij}/n$, het $j^e$-kolomgemiddelde.

$\bar{y} := \bar{y}_{\cdot}u$, een vector met alle componenten gelijk aan $\bar{y}_{\cdot} = \sum y_i/n$.

$P := \frac{1}{n}U(= \frac{1}{n}uu')$, de orthogonale projector op $u$ (zie 1.91);

$\bar{X}_1 := PX_1$. Er blijkt dat $(\bar{X}_1)_{i*} = \bar{x}$ voor alle $i$.

$X_m := X_1 - \bar{X}_1 = P^{\perp}X_1$ d.w.z. trek van elke kolom $x_{*j}$ van $X_1$ de vector $\bar{x}_{\cdot j}u$, het kolomgemiddelde, af.

$y_m := y - \bar{y} = (I - P)y = P^{\perp}y$.

Hieruit volgt direct:

5.7. $u^2 = u'u = n$; $u'y = n\bar{y}_.$; $\bar{X}_1 = u\bar{x}'$; $u'X_1 = n\bar{x}'$;

In gepartitioneerde vorm worden de normaal vergelijkingen (5.3):

$$\begin{pmatrix} n & n\bar{x}' \\ n\bar{x} & X_1'X_1 \end{pmatrix}\begin{pmatrix} b_0 \\ b_m \end{pmatrix} = \begin{pmatrix} u'y \\ X_1'y \end{pmatrix}$$

en $y = X\beta$ wordt:

$$y = (u \;\; X_1)\begin{pmatrix} \beta_0 \\ \beta_m \end{pmatrix} = u\beta_0 + X_1\beta_m.$$

Dus $nb_0 + n\bar{x}'b_m = u'y = n\bar{y}$.

5.8. $$\underline{b}_0 = \bar{\underline{y}}_. - \bar{x}'\underline{b}_m$$

Verder is $y - \bar{y} = P^{\perp}y = P^{\perp}u\beta_0 + P^{\perp}X_1\beta_m$

ofwel $y_m = X_m\beta_m$ nl. $P^{\perp}u = 0$.

Dus analoog aan (5.4): $\underline{b}_m = (X_m'X_m)^{-1}X_m'\underline{y}_m$.

<u>Opmerking</u>: $X_m'X_m$ is inderdaad regulier nl. $X$, dus $X_1$ is van volle kolomrang. $X_m$ ontstaat door van $X_1$ een veelvoud van 1^e kolom van $X$ af te trekken, waardoor de rang niet verandert, dus $X_m$ blijft van volle kolomrang, en daarmede is $X_m'X_m$ regulier, zodat $(X_m'X_m)^{-1}$ bestaat.

f. <u>Verschillende schattingsmethoden voor $\beta$</u>

$\beta$ kan op verschillende manieren geschat worden:

i) De <u>kleinste kwadratenmethode</u> houdt het minimaliseren in van $\| \underline{y} - X\beta \|^2$ met betrekking tot $\beta$ en geeft $\underline{b} = (X'X)^{-1}X'\underline{y}$.

ii) De <u>gewogen kleinste kwadratenmethode</u> houdt het minimaliseren in van $\| \Sigma^{-T}(y - X\beta) \|^2 = (y - X\beta)'\Sigma^{-2}(y - X\beta)$ met betrekking tot $\beta$ als var $\underline{e} = \Sigma^2$ en geeft als schatter voor $\beta$ : $\underline{\tilde{b}} = (X'\Sigma^{-2}X)^{-1}X'\Sigma^{-2}\underline{y}$.
Is $\Sigma^2 = \sigma^2 I$ dan wordt $\underline{\tilde{b}} = \underline{b}$. Deze $\underline{\tilde{b}}$ volgt direct uit i) door $\underline{y}$ te vervangen door $\Sigma^{-T}\underline{y}$ en X door $\Sigma^{-T}X$!

iii) <u>Maximum likelihoodmethode</u>. Deze houdt een veronderstelling in over de verdeling van $\underline{e}$. Hiervoor wordt vaak de multinormale verdeling genomen d.w.z. $\underline{e} \sim N_n(0,\Sigma^2)$ en dus $\underline{y} \sim N_n(X\beta,\Sigma^2)$.
Het maximaliseren van de likelihoodfunctie

$$L = (2\pi)^{-n/2}|\Sigma|^{-1}\exp[-\tfrac{1}{2}(y - X\beta)'\Sigma^{-2}(y - X\beta)]$$

met betrekking tot $\beta$, geeft de M.L. schatter voor $\beta$ die juist $\underline{\tilde{b}}$ blijkt te zijn. (Dus de M.L. schatter voor $\beta$ is onder aanname van de multinormale verdeling voor $\underline{e}$, de gewogen kleinste kwadraten schatter).

iv) <u>De beste lineaire zuivere schatter</u> (<u>blue</u> = best lineair unbiased estimator). Als var $\underline{e} = \Sigma^2$, wordt gezocht naar de beste lineaire zuivere schatter voor $t'\beta$ ($t$ willekeurige vector). Onder de voorwaarde dat zo'n schatter van het type $\alpha'\underline{y}$ (lineair) is en geldt $\mathcal{E}\alpha'\underline{y} = t\beta$ blijkt $t'\underline{\tilde{b}}$ de kleinste variantie te hebben en dus de "beste" lineaire zuivere schatter te zijn voor $t'\beta$. De keuze van $t$ is willekeurig.
Neem $t = \delta_{i*}$ dan is dus $t'\underline{\tilde{b}} = \underline{\tilde{b}}_i$ oftewel $\underline{\tilde{b}}_i$ is blue voor $t'\beta = \beta_i$. Aangezien dit geldt voor elke i, noemen we $\underline{\tilde{b}}$ blue voor $\beta$.

g. <u>Eigenschappen van de schatters</u>

In het vervolg wordt ondersteld var $\underline{e} = \sigma^2 I$.

i) <u>Zuiverheid</u>

$$\mathcal{E}\underline{b} = \mathcal{E}[(X'X)^{-1}X'\underline{y}] = (X'X)^{-1}X'X\beta = \beta,$$

dus ook $\mathcal{E}\underline{b}_m = \beta_m$.

ii) <u>Variantie</u>

$$\text{var } \underline{b} = (X'X)^{-1}X'(\sigma^2 I)X(X'X)^{-1} = (X'X)^{-1}\sigma^2 \quad \text{zie (2.38).}$$

Analoog var $\underline{b}_m = (X_m'X_m)^{-1}\sigma^2$.

Nu is $\text{cov}(\bar{\underline{y}},\underline{b}_m) = \text{cov}(P\underline{y},AP^{\perp}\underline{y})$ met $A := (X_m'X_m)^{-1}X_m' \Rightarrow P\sigma^2 IP^{\perp}A' = 0$.

5.10. Gevolg: $\text{cov}(\underline{b}_0,\underline{b}_m) = \text{cov}(\bar{\underline{y}} - \bar{x}'\underline{b}_m,\underline{b}_m) = -\bar{x}'(X_m'X_m)^{-1}\sigma^2$ en

$$\text{var } \underline{b}_0 = \text{var}(\bar{\underline{y}}_{\cdot} - \bar{x}'\underline{b}_m) = \text{var } \bar{\underline{y}}_{\cdot} + \bar{x}'\text{var } \underline{b}_m\bar{x} =$$

$$= \sigma^2[1/n + \bar{x}'(X_m'X_m)^{-1}\bar{x}].$$

iii) Schatter voor $\mathcal{E}\underline{y} = X\beta$

5.11. Deze schatter is $\hat{\underline{y}} = X\underline{b} = X(X'X)^{-1}X'\underline{y} = P_X\underline{y}$. De waarden $\hat{y}$ zijn de geschatte verwachtingswaarden bij de waargenomen y's (ook wel voorspelde of geschatte y-waarden genoemd).

Evident is $\mathcal{E}\hat{\underline{y}} = X\mathcal{E}\underline{b} = X\beta$ en

5.12. $$\text{var } \hat{\underline{y}} = \text{var}(P_X\underline{y}) = P_X\sigma^2 = X(X'X)^{-1}X'\sigma^2 .$$

Is nu $\xi' := (\xi_0,\xi_1,\ldots,\xi_k)$ een set x-waarden, een andere dan die voor de bepaling van $\underline{b}$ gebruikt wordt.

We schatten dan $\xi'\beta$ door $\xi'\underline{b}$ en $\text{var}(\xi'\underline{b}) = \xi'(X'X)^{-1}\xi\sigma^2$; Zij $\xi' = (\xi_0,\xi_m') = (1,\xi_m')$.

Dan wordt met $(X'X)^{-1}$ als onder 5.7 gepartitioneerd

$$\text{var}(\xi'\underline{b}) = \sigma^2[1/n + (\xi_m-\bar{x})'(X_m'X_m)^{-1}(\xi_m-\bar{x})] .$$

We zien hieruit dat de variantie groter wordt naarmate $\xi$ meer van $\bar{x}$ afwijkt. Beschouw nu een "toekomstige" (a.h.w. nà de bepaling van $\underline{b}$) set x-waarden $x_f$. Dan is $\underline{y}_f = x_f'\beta + \underline{e}_f$ en wordt de schatter van $\mathcal{E}\underline{y}_f$ gegeven door $\underline{x}_f'\underline{b}$. Omdat $\underline{e}_f$ niet kan worden geschat of waargenomen, is de beste beschikbare voorspelling voor $\underline{y}_f$ ook $x_f'\underline{b} =: \tilde{\underline{y}}_f$.

Van belang is dan te weten in hoeverre $\underline{y}_f$ varieert om zijn voorspelde waarde $\tilde{\underline{y}}_f = x_f'\underline{b}$. Berekenen we $\text{var}(\underline{y}_f - \tilde{\underline{y}}_f) = \text{var } \underline{y}_f + \text{var } \tilde{\underline{y}}_f$ op grond van het feit dat $\underline{y}_f$ onafhankelijk is van $\underline{y}_1,\ldots,\underline{y}_n$ waaruit $\underline{b}$ bepaald is.

Nu is $\text{var}(\underline{y}_f - \tilde{\underline{y}}_f) = [x_f'(X'X)^{-1}x_f + 1]\sigma^2$ nl. $\text{var } \underline{y}_f = \text{var } \underline{e}_f = \sigma^2$.

h. Restkwadratensom

De restkwadratensom (SSE = Sum of Squares of Error) wordt gedefinieerd door:

5.13. $$\mathrm{SSE} := \|P_X^{\perp}y\|^2 = \|y - \hat{y}\|^2 = \sum_1^n (y_i - \hat{y}_i)^2 .$$

Dit wordt:

$$\mathrm{SSE} = y'P_X^{\perp}y = y'(I - P_X)y = y'y - (P_Xy)'y = y'y - b'X'y .$$

Deze gaan we uitdrukken in grootheden, gecorrigeerd voor het gemiddelde.

$$\hat{y}_m := P_{X_m} y_m .$$

5.14. Dus

$$\mathrm{SSE}_m := \|y_m - \hat{y}_m\|^2 = y_m'y_m - b_m'X_m'y_m,$$

analoog aan (5.13).

Nu is $P_{X_m} = P_X - P$ (Ga na) dus $\hat{y}_m = (P_X - P)y = \hat{y} - \bar{y}$.

Dus

$$y - \hat{y} = y - \bar{y} + \bar{y} - \hat{y} = y_m - \hat{y}_m,$$

oftewel

$$\mathrm{SSE} = \mathrm{SSE}_m$$

(zie ook fig. 5.1).

Opmerking: Er geldt: de som der residuen $\Sigma(y_i - \hat{y}_i) = 0$ nl.

$$u'(y - \hat{y}) = (P_Xu)'(P_X^{\perp}y) = u'P_XP_X^{\perp}y = 0 \quad (u \in \langle X\rangle) .$$

k. Schatter voor $\sigma^2$

$\mathcal{E}(\mathrm{SSE}) = \mathcal{E}\|P_X^{\perp}y\|^2 = \sigma^2 r(P_X^{\perp}) = \sigma^2(n - r)$ als $r = r(X) = k + 1$.

Dit volgt uit 2.36 en $P_X^{\perp}(X\beta) = 0$.

5.15. Een zuivere schatter voor $\sigma^2$ is dus $\underline{s}^2 := \mathrm{SSE}/(n - r)$.

1. Splitsing van de totale kwadratensom

Daar $y - \hat{y} = P_X^{\perp}y$ en $\hat{y} = P_Xy$, is $y - \hat{y} \perp \hat{y}$ en geldt: $y = \hat{y} + (y - \hat{y})$, dus

$\|y\|^2 = \|\hat{y}\|^2 + \|y - \hat{y}\|^2$ oftewel SST = SSR + SSE hetgeen betekent: Sum of Squares for Total = Sum of Squares due to Regression + Sum of Squares for Error.

5.16. Dus SSR = $\hat{y}'\hat{y} = b'X'Xb = b'X'y$ en tevens $= y^2 - y_m^2 + b_m'X_m'y_m$ (zie 5.14). Daar $y - \bar{y} = P^{\perp}y$ en $\bar{y} = Py$, is $y - \bar{y} \perp \bar{y}$ en geldt $y = y - \bar{y} + \bar{y}$, dus $\|y\|^2 = \|y - \bar{y}\|^2 + \|\bar{y}\|^2 = \|y_m\|^2 + \|\bar{y}\|^2$ oftewel SST = $SST_m$ + SSM hetgeen betekent: SST = Sum of Squares for Total, corrected for mean + Sum of Squares due to the mean.

Aangezien het voor de "kwaliteit" van de aanpasssing van het model geen verschil maakt of uitgegaan wordt van $y$ of $y_m$ (de schatter $\underline{b}_m$ is in beide gevallen nl. hetzelfde daar $X_m'y = X_m'y_m$), moet de restkwadratensom SSE waaruit $s^2$ wordt afgeleid in beide gevallen dezelfde zijn, SSE = $SSE_m$. Definiëren we nog $SSR_m$ = SSR - SSM oftewel Sum of Squares due to Regression, corrected for the mean.

Zo ontstaat analoog de splitsing: $SST_m = SSR_m + SSE_m$, nl.:

$$y_m = (\hat{y} - \bar{y}) + (y - \hat{y}) = \hat{y}_m + (y_m - \hat{y}_m).$$

5.17. Uit 5.16 volgt $SSR_m = b_m'X_m'y_m$; $SST_m = y_m^2 = y^2 - n\bar{y}_.^2$;
Uit de volgende figuur blijkt de splitsing van de totale kwadratensom duidelijk; Bij de getekende vectoren zijn hun lengten in het kwadraat gegeven.

We lezen hier uit af:

$SSR_m$ = SSR - SSM
$SST_m$ = SST - SSM
$SST_m = SSR_m + SSE_m$
SST = SSR + SSE en
SST = SSM + $SSR_m$ + SSE.
$SSE_m$ = SSE

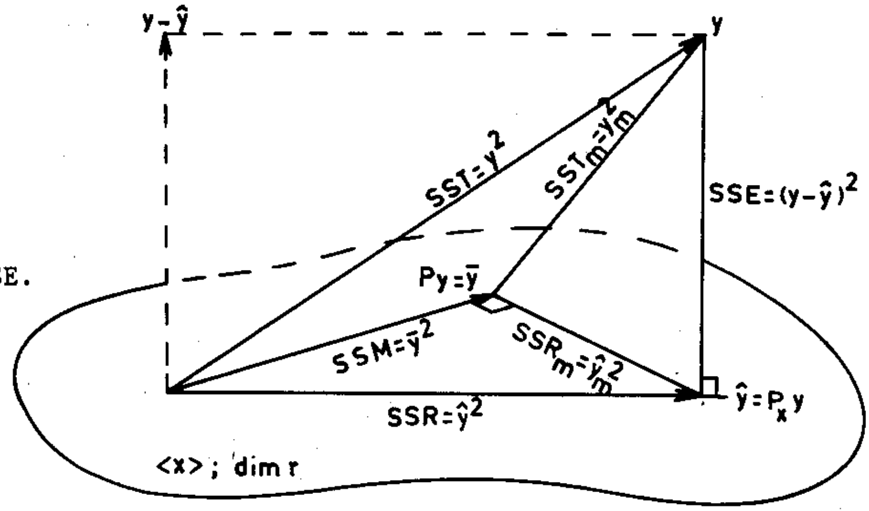


fig. 5.1.

<u>Opmerking</u>: We zagen $P_X - P = P_{X_m}$. Dus

5.18. $$SSR = \| P_X y \|^2 \text{ en } SSR_m = \| P_{X_m} y_m \|^2 = \| P_{X_m} y \|^2$$

Er geldt: $PP_{X_m} = P_{X_m} P = 0$.

## § 6. Verdelingen in het regressiemodel

In het vervolg nemen we steeds aan $\underline{e} \sim N_n(0,\sigma^2 I)$.
Daar $\underline{y} = X\beta + \underline{e} =: \eta + \underline{e}$, volgt hieruit direct

6.1. $$\underline{y} \sim N_n(\eta,\sigma^2 I) \ .$$

Nu was $\underline{b} = (X'X)^{-1}X'\underline{y}$ en $r[(X'X)^{-1}X'] = r(X') = r = k+1$ dus met 3.3 en het feit dat $\underline{b}$ zuiver is, geldt:

6.2. $\underline{b} \sim N_{k+1}(\beta, X^{-2}\sigma^2)$, analoog $\underline{b}_m \sim N_k(\beta_m, X_m^{-2}\sigma^2)$.

6.3. Stelling: $SSE \simeq \sigma^2 \chi^2_{n-r}$.

Bewijs: $SSE = \| P_X^{\perp}\underline{y} \|^2$ en met 4.12 is dus $SSE/\sigma^2 \simeq \chi^2_{n-r}$ nl. $r(P_X^{\perp}) = r(I - P_X) =$ $= n - r$; De niet-centraliteitsparameter $\delta = 0$ daar

$$\sigma^2\delta = \| P_X^{\perp}\, \mathcal{E}\underline{y} \|^2 = \| P_X^{\perp}\, X\beta \|^2 = 0 \ .$$

Gevolg: $(n-r)\underline{s}^2/\sigma^2 \simeq \chi^2_{n-r}$ (zie 5.15).

6.4. Stelling: $\underline{b}$ onafhankelijk $\underline{s}^2$.

Bewijs: Dit volgt direct uit 4.13 nl. $\underline{b} = (X'X)^{-1}X'\underline{y} =: B\underline{y}$ en

$$SSE = \| P_X^{\perp}\underline{y} \|^2 = \underline{y}'P_X^{\perp}\underline{y} =: \underline{y}'A\underline{y} \ .$$

Nu is $A\Sigma^2 B' = P_X^{\perp} X X^{-2}\sigma^2 = 0$ $(= B\Sigma^2 A)$. Verder is $SSE = (n-r)\underline{s}^2$.

### a. Niet-centrale $\chi^2$-verdelingen

In § 5 ℓ is aangegeven hoe de totale kwadratensom kan worden gesplitst. We gaan eens na hoe deze kwadratensommmen verdeeld zijn.
De verdelingen volgen alle direct uit 4.12 nl.:

6.5. $SSR := \| P_X y \|^2 \simeq \sigma^2 \chi_r^2(\delta) \qquad r = r(P_X) \qquad \delta = \| P_X \eta \|^2/\sigma^2 = \| \eta \|^2/\sigma^2$

$SSR_m := \| P_{X_m} y \|^2 \simeq \sigma^2 \chi_{r-1}^2(\delta_m) \qquad r - 1 = r(P_{X_m}) \qquad \delta_m = \| \eta_m \|^2/\sigma^2$

$\eta_m := X_m \beta_m$

$SST := \| I y \|^2 \simeq \sigma^2 \chi_n^2(\delta) \qquad n = r(I)$

$SST_m := \| P^\perp y \|^2 \simeq \sigma^2 \chi_{n-1}^2(\delta_m) \qquad n - 1 = r(P^\perp)$

$SSM := \| P y \|^2 \simeq \sigma^2 \chi^2(\lambda) \qquad 1 = r(P) \qquad \lambda = (\mathcal{E}\bar{y})^2/\sigma^2 = \| P\mathcal{E} y \|^2/\sigma^2$

$SSE := \| P_X^\perp y \|^2 \simeq \sigma^2 \chi_{n-r}^2 \qquad n - r = r(P_X^\perp) \qquad \| P_X^\perp \mathcal{E} y \|^2 = 0 .$

Daar $P_X P_X^\perp = 0$, $PP_X^\perp = 0$, $PP_{X_m} = 0$ en $P_{X_m} P_X^\perp = 0$ volgt met 4.13* dat resp. SSR en SSE, SSM en SSE, SSM en $SSR_m$, $SSR_m$ en SSE alle onafhankelijk zijn. Dit alles lezen we ook direct af uit fig. 5.1. Hierbij wordt de rang = dimensie van de ruimte waarop geprojecteerd wordt.

b. F-verdelingen

Delen we de kwadratensom SS door bijbehorend aantal vrijheidsgraden $\nu$, dan krijgen we de MS (mean Square). Zo wordt:

$$MSM := SSM/1 \; ; \quad MSR_m = SSR_m/(r-1) \; ; \; MSE := SSE/(n-r) \; ;$$

$$MSR := SSR/r \; .$$

Quotiënten van deze onafhankelijke grootheden geven steeds een niet-centrale F-variabele.
Zo is:

6.6. $F(R) := MSR/MSE \sim F_{n-r}^{r}(\delta) \; ; \quad F(R_m) := MSR_m/MSE \sim F_{n-r}^{r-1}(\delta_m) \; ;$

$F(M) := MSM/MSE \sim F_{n-r}^{1}(\lambda) \; .$

c. Variantie-analyse

Het geheel van de berekeningen, vrijheidsgraden e.d. kunnen we overzichtelijk weergeven in een zogenaamde variantie-analysetabel.
Zo'n tabel geeft de bronnen van variatie, SS, aantal vrijheidsgraden $\nu$, MS, en de EMS (Expectation of Mean Square). Volledigheidshalve zijn de verdelingen der SS ook opgegeven, hoewel deze in zo'n tabel niet thuishoren.

Model: $\underline{y} = X\beta + \underline{e}$ ; $\underline{e} \sim N_n(0,\sigma^2 I)$ ; $r(X) = r = k + 1$.

| Bron | SS | | $\nu$ | verdeling SS | MS = SS/$\nu$ | EMS |
|---|---|---|---|---|---|---|
| SSR | $\Vert P_X y\Vert^2 = \Vert\hat{y}\Vert^2$ | $b'X'y$ | r | $\sigma^2\chi^2_\nu(\delta)$ | | $\sigma^2 + \sigma^2\delta/\nu$ |
| SSE | $\Vert P_X^\perp y\Vert^2 = \Vert y - \hat{y}\Vert^2$ | $\underline{y^2} - b'X'y$ | n-r | $\sigma^2\chi^2_\nu$ | | $\sigma^2$ |
| SST | $\Vert Iy\Vert^2 = \Vert y\Vert^2$ | $\underline{y^2}$ | n | $\sigma^2\chi^2_\nu(\delta)$ | | |
| | Gecorrigeerd voor het gemiddelde levert: | | | | | |
| SSM | $\Vert Py\Vert^2 = \Vert\bar{y}\Vert^2$ | $n\bar{y}^2_{\cdot}$ | 1 | $\sigma^2\chi^2_\nu(\lambda)$ | | $\sigma^2 + \sigma^2\lambda/\nu$ |
| $SSR_m$ | $\Vert P_{X_m} y\Vert^2 = \Vert\hat{y}_m\Vert^2$ | $b_m'X_m'y_m$ | r-1 | $\sigma^2\chi^2_\nu(\delta_m)$ | | $\sigma^2 + \sigma^2\delta_m/\nu$ |
| SSE | $\Vert P_X^\perp y\Vert^2 = \Vert y - \hat{y}\Vert^2$ | $\underline{y^2_m} - b_m'X_m'y_m$ | n-r | $\sigma^2\chi^2_\nu$ | | $\sigma^2$ |
| $SST_m$ | $\Vert P^\perp y\Vert^2 = \Vert y_m\Vert^2$ | $\underline{y^2_m}$ | n-1 | $\sigma^2\chi^2_\nu(\delta_m)$ | | |

tabel 6.1 en 6.2

Opmerking:

i) Evenals $P_X y = \hat{y} = Xb$, is $P_{X_m} y =: \hat{y}_m = X_m b_m$.

$y_m - \hat{y}_m = (y - \bar{y}) - (\hat{y} - \bar{y}) = y - \hat{y}$ dus $SSE_m = SSE$.

$\underline{y^2_m} = \underline{y^2} - n\bar{y}^2_{\cdot}$;

ii) $\mathcal{E}MS = \mathcal{E}SS/\nu = \mathcal{E}[\sigma^2\chi^2(\delta)/\nu] = \sigma^2(\nu + \delta)/\nu = \sigma^2 + \sigma^2\delta/\nu$ (zie 4.2).

d. Het toetsen van hypothesen

De grootheden F(R), $F(R_m)$ en F(M) zijn voor bepaalde hypothesen geschikte toetsingsgrootheden. Onder de bijbehorende nulhypothese zijn deze centraal verdeeld. Zo is (zie ook 6.6).

$F(R) \sim F^r_{n-r}(\delta)$, $H_0$ : $\delta = 0$ dus $X\beta = 0$ dus $\beta = 0$ (nl. X volle kolomrang)

$F(R_m) \sim F^{r-1}_{n-r}(\delta_m)$, $H_0$ : $\delta_m = 0$ dus $X_m\beta_m = 0$ dus $\beta_m = 0$ (nl. $X_m$ volle kolomrang)

$F(M) \sim F^1_{n-r}(\lambda)$, $H_0$ : $\lambda = 0$ oftewel $\mathcal{E}\bar{y} = 0$ .

Deze geven aanleiding tot rechtseenzijdige toetsen. Immers onder $H_1 : \beta \neq 0$ geldt bv. $F(R) \sim F^{r}_{n-r}(\delta)$ en $P[F(R) \geq F(\alpha) \mid H_1] > P[F(R) \geq F(\alpha) \mid H_0]$ hetgeen ondermeer volgt uit het feit dat de verwachtingswaarde van F(R) groter is onder $H_1$ dan onder $H_0$ (nl. $\mathcal{E}\underline{F}^{\nu_1}_{\nu_2} = \nu_2(1 + \delta/\nu_1)/(\nu_2 - 2)$ zie Searle en $\delta > 0$).

Als een "erg grote" waarde van F(R) gevonden wordt, is er dus reden om $H_0$ te verwerpen, omdat grote waarden waarschijnlijker zijn onder $H_1$ dan onder $H_0$.

Opmerking: De F-toets is een likelihood-ratiotoets (Searle pag. 124 e.v.).

De grootheid die het meest gebruikt wordt als toetsingsgrootheid is $F(R_m)$. Onder de nulhypothese $H_0 : \beta_m = 0$ heeft $F(R_m)$ een centrale F-verdeling. Bij het toetsen van $H_0$ tegen $H_1 : \beta_m \neq 0$ wordt $H_0$ weer verworpen als $F(R_m) \geq F(\alpha)$. Als $F(R_m)$ een significante waarde heeft, wordt besloten dat niet alle componenten van $\beta_m$ nul zijn. Dat wil niet zeggen dat ze allemaal van nul verschillen. De volgende stap is dan het uitzoeken welke componenten van nul verschillen, om te kunnen concluderen welke variabele wel en welke niet bijdragen tot het verklaren van de totale variatie in de y-waarnemingen. Of er meer verklarende variabelen nodig zijn om de totale variatie te verklaren kan onder andere getoetst worden door er meerdere in een tweede analyse op te nemen en te onderzoeken of ze een rol spelen. Een andere mogelijkheid is als we per combinatie van x-waarden meerdere y-waarnemingen hebben. De restkwadraatsom kan dan in twee stukken worden gesplitst, één binnen de subgroepen met dezelfde x-waarden (SS for Pure Error, SSPE) en de rest (SS of Lack of Fit, SSLF). Als er p verschillende x-combinaties en in het totaal n waarnemingen zijn, heeft $\frac{SSLF/(p - r)}{SSPE/(n - p)}$ een centrale $F^{p-r}_{n-p}$-verdeling, onder de hypothese dat het model juist is.

e. Betrouwbaarheidsintervallen

We weten $\underline{b} \sim N_{k+1}(\beta, X^{-2}\sigma^2)$ en dus $\underline{b}_i \sim N(\beta_i, X^{-2}_{ii}\sigma^2)$, $i = 1,\ldots,k$. Nu is

$$\frac{(\underline{b}_i - \beta_i)/\sqrt{X^{-2}_{ii}\sigma^2}}{\sqrt{\underline{s}^2/\sigma^2}} \simeq \frac{\underline{\chi}}{\sqrt{\chi^2_\nu/\nu}} \simeq \underline{t}_\nu$$

met $\nu = n - r$, daar $\underline{b}$ onafhankelijk $\underline{s}^2$ is.

Hieruit volgt een betrouwbaarheidsinterval voor $\beta_i$ met betrouwbaarheid $1 - \alpha$.

6.7. $$b_i - t_\nu(\tfrac{1}{2}\alpha)s\sqrt{X^{-2}_{ii}} < \beta_i < b_i + t_\nu(\tfrac{1}{2}\alpha)s\sqrt{X^{-2}_{ii}} \; ; \qquad \nu = n - r \; .$$

Het toetsen van de nulhypothese $H_0 : \beta_i = 0$ kan eenvoudig worden uitgevoerd door te onderzoeken of 6.7 de waarde $\beta_i = 0$ bevat.

Omdat de $\underline{b}_i$'s afhankelijk kunnen zijn en omdat we dezelfde s gebruiken is het niet geoorloofd een simultaan betrouwbaarheidsinterval voor bijvoorbeeld $\beta_i$ en $\beta_j$ op te stellen door combinatie van twee intervallen als in 6.7 gegeven, tot een rechthoek in het $(\beta_i,\beta_j)$-vlak. We gaan dan anders te werk:

$$\| \underline{\hat{y}} - \eta \|^2 = \| X(\underline{b} - \beta) \|^2 \sim \sigma^2 \chi^2_r$$

nl. met 4.11 $X^{\underline{2}}/\sigma^2 \; X^{-\underline{2}}\sigma^2 = I$ (idempotent), $r(X) = r$ en $\mathcal{E}X(\underline{b} - \beta) = 0$.
Dus

6.8. $$\frac{\| \underline{\hat{y}} - \eta \|^2 / r}{\| \underline{\hat{y}} - \underline{y} \|^2 / (n - r)} = \frac{\| X(\underline{b} - \beta) \|^2 / r}{\underline{s}^2} \simeq \underline{F}^r_{n-r} \; ,$$

teller en noemer onafhankelijk daar $\underline{b}$ en $\underline{s}^2$ onafhankelijk zijn, zie (6.4).
De ongelijkheid $\| X(\underline{b} - \beta) \|^2 \leq rs^2F(\alpha)$ geeft een betrouwbaarheidsellipsoïde voor $\beta$ in de parameterruimte. Voor $k = 1$ is dit een ellips.
Het is een ellipsoïde omdat ontstaat $(b - \beta)'X^{\underline{2}}(b - \beta) \leq c$ en $X^{\underline{2}} > 0$, dus alle $\lambda_i > 0$.

In 5.12 is als schatter voor $\xi'\beta$ de grootheid $\xi'\underline{b}$ genomen.
De betrouwbaarheidsgrenzen worden voor $\xi'\beta$

$$\xi'b \pm t_{n-r}(\tfrac{1}{2}\alpha)s\sqrt{\xi'X^{-\underline{2}}\xi} \; .$$

Voor $k = 1$, dus $r = 2$, wordt dit

$$b_0 + b_1\xi_m \pm t_{n-2}(\tfrac{1}{2}\alpha)s\sqrt{1/n + (\bar{x} - \xi_m)^2/\| x - \bar{x} \|^2} \; .$$

## § 7. De algemene lineaire hypothese

### a. Het toetsen van lineaire hypothesen

In de literatuur over lineaire modellen worden verschillende hypothesen besproken, o.a.

i) $H_0 : \beta = 0$

ii) $H_0 : \beta = \beta^*$

iii) $H_0 : \alpha'\beta = \gamma$

iv) $H_0 : \beta^{(1)} = 0$

d.w.z. $\beta' = (\beta^{(1)'}\beta^{(2)'})$ en $\beta^{(1)}$ heeft bv. q componenten.
Al deze lineaire hypothesen kunnen zoals we zullen zien onder c), als speciale gevallen beschouwd worden van een algemene lineaire hypothese:

Beschouw de algemene lineaire hypothese $H_0 : K\beta = \gamma$ met $K \in \mathfrak{M}^q_{q,k+1}$ dus van volle rijrang; $\gamma$ is een vector van gespecificeerde constanten ($q \leq r$ zie 10.2).
Om $H_0 : K\beta = \gamma$ te kunnen toetsen hebben we een toetsingsgrootheid nodig.

Nu is $K\underline{b} - \gamma \sim N_q(K\beta - \gamma, V^{-2}\sigma^2)$ met $V^{-2} := \mathrm{var}(K\underline{b})/\sigma^2 = KX^{-2}K' > 0$ (volgens 1.66) zodat deze notatie geoorloofd is.
Definieer $SSQ := \| V(K\underline{b} - \gamma)\|^2$. Nu geldt:

7.1. $SSQ \simeq \sigma^2\chi^2_q(\delta_Q)$ nl. met 4.12 $V^2/\sigma^2\, V^{-2}\sigma^2 = I$ idempotent en $r(V) = r(V^2) = q$ (nl. $V^2_{q\times q}$ en regulier!), $\delta_Q = \| V(K\beta - \gamma)\|^2/\sigma^2$.

Nu is SSQ onafhankelijk SSE daar $\underline{b}$ onafhankelijk $\underline{s}^2$ is. Vandaar dat

$$F(Q) := \frac{MSQ}{MSE} = \frac{SSQ/q}{\underline{s}^2} \simeq F^q_{n-r}(\delta_Q) .$$

Onder de nulhypothese $H_0 : K\beta = \gamma$ is $\delta_Q = 0$ en is $F(Q)$ centraal verdeeld.

### b. Schatten van $\beta$ onder $H_0$: $K\beta = \gamma$

We vinden de schatter $\underline{b}_{X_H}$ door $\| \underline{y} - X\beta \|^2$ te minimaliseren naar $\beta$ onder de restricties $K\beta = \gamma$.
Beschouw $f(\beta,\theta) = (y - X\beta)'(y - X\beta) + 2\theta'(K\beta - \gamma)$.

7.2. $$\partial f/\partial\beta' = -2X'y + 2X'X\beta + 2K'\theta \ .$$

7.3. $$\partial f/\partial\theta' = 2(K\beta - \gamma) \ .$$

Nul stellen en oplossen geeft:

7.4. Uit (7.2) $\underline{b}_{X_H} = (X'X)^{-1}(X'\underline{y} - K'\theta_H) = \underline{b} - X^{\underline{-2}}K'\theta_H$ .

Uit 7.3 en 7.4 volgt: $K\underline{b}_{X_H} = K\underline{b} - KX^{\underline{-2}}K'\theta_H = \gamma$, oftewel $\theta_H = V^{\underline{2}}(K\underline{b} - \gamma)$ zodat

7.5. $$\underline{b}_{X_H} = \underline{b} - X^{\underline{-2}}K'V^{\underline{2}}(K\underline{b} - \gamma) \ .$$

De restkwadraatsom onder $H : K\beta = \gamma$ wordt dan $SSE_Q = \|y - \hat{y}_{X_H}\|^2 := \|y - Xb_{X_H}\|^2 =$ $= \|y - \hat{y}\|^2 + \|\hat{y} - \hat{y}_{X_H}\|^2$ daar $y - \hat{y} \perp \langle X\rangle$ en $\hat{y} - \hat{y}_{X_H} = X(b - b_{X_H}) \in \langle X\rangle$. Dit is juist $= SSE + SSQ$ nl.:

$$\|\hat{y} - \hat{y}_{X_H}\|^2 = \|X(b - b_{X_H})\|^2 = (Kb - \gamma)'V^{\underline{2}}KX^{\underline{-2}}X'XX^{\underline{-2}}K'V^{\underline{2}}(Kb - \gamma) =$$

$$= \|V(Kb - \gamma)\|^2 = SSQ \ .$$

c. <u>Speciale lineaire gevallen</u>

Bekijken we eens de speciale gevallen genoemd onder a).

i) $H_0 : \beta = 0$.
Neem $K = I$ en $\gamma = 0$ dus $q = r$, $V^{\underline{2}} = X^{\underline{2}}$ en $SSQ = SSR$ dus $F(Q) = F(R)$.
$b_{X_H} = b - b = 0$ hetgeen nogal logisch is, omdat onder $H_0$ geldt $\beta = 0$.

ii) $H_0 : \beta = \beta^*$.
Neem $K = I$ en $\gamma = \beta^*$ dus $q = r$, $V^{\underline{2}} = X^{\underline{2}}$, dan is

$$F(Q) = \|X(\underline{b} - \beta^*)\|^2/r\underline{s}^2 \simeq F^r_{n-r}, \text{ met } \delta_\varphi = \|X(\beta - \beta^*)\|^2/\sigma^2 \ (= 0 \text{ onder } H_0).$$

Vanzelfsprekend wordt $b_{X_H} = b - X^{\underline{-2}}V^{\underline{2}}(b - \beta^*) = \beta^*$.

iii) $H_0 : \alpha'\beta = \gamma$.
Neem $K = \alpha'$, $\gamma = \gamma$ dus $q = 1$ dan wordt

$$F(Q) = SSQ/\underline{s}^2 = \frac{(\alpha'\underline{b} - \gamma)^2}{\alpha'X^{\underline{-2}}\alpha\underline{s}^2} \simeq \underline{F}^1_{n-r} \simeq \underline{t}^2_{n-r} \ (\text{onder } H_0) \ .$$

$$\underline{b}_{X_H} = \underline{b} - X^{-2}\alpha[\alpha' X^{-2}\alpha]^{-1}(\alpha' b - \gamma) = \underline{b} - \frac{\alpha' b - \gamma}{\alpha' X^{-2}\alpha} X^{-2}\alpha \ .$$

iv) $H_0 : \beta^{(1)} = 0$ d.w.z. $\beta_i = 0$, $i = 0,\ldots,q-1$ $(q < k + 1)$.
Zonder verlies van algemeenheid kunnen we de eerste q componenten nul nemen.
Neem $K = (I_q \ \ 0)$, $\gamma = 0$ (dus inderdaad K van volle rijrang).
We partitioneren als volgt:

$$\beta' = (\beta^{(1)'}\beta^{(2)'}) \ ; \ \ \underline{b}' = (\underline{b}^{(1)'}\underline{b}^{(2)'}),\ X^{-\underline{2}} = \begin{pmatrix} T_{qq} & T_{q,r-q} \\ T_{r-q,q} & T_{r-q,r-q} \end{pmatrix}$$

Dus $V^{-\underline{2}} = KX^{-\underline{2}}K' = T_{qq}$.

$$F(Q) = SSQ/q\underline{s}^2 = \underline{b}^{(1)'} T_{qq}^{-1} \underline{b}^{(1)}/q\underline{s}^2 \simeq \underline{F}^q_{n-r} \text{ (centraal onder } H_0) \ .$$

$$\underline{b}_{X_H} = \underline{b} - X^{-2}\begin{pmatrix} I_q \\ 0 \end{pmatrix} T_{qq}^{-1}(\underline{b}^{(1)} - 0) = \begin{pmatrix} 0 \\ b^{(2)} - T_{r-q,q}T_{qq}^{-1}b^{(1)} \end{pmatrix}$$

(inderdaad $b^{(1)} = 0$).

d. Gereduceerde modellen

Het model $\underline{y} = X\beta + \underline{e}$ zonder enige restrictie opgelegd aan de parameters $\beta$ heet wel het volle model in tegenstelling tot het model $\underline{y} = X\beta + \underline{e}$ met de restricties $K\beta = \gamma$ wat het gereduceerde model genoemd wordt. In het volle model is de restkwadraatsom gelijk aan SSE, in het gereduceerde model aan $SSE_Q = SSE + SSQ > SSE$ daar $SSQ > 0$. Dit is evident nl. het minimum van $\| y - X\beta \|^2$ naar $\beta$ zonder enige restricties is $\leq$ minimum onder de restricties $K\beta = \gamma$.
Omdat geldt

SSQ = reskwadraatsom (gered. model) - restkwadraatsom (volle model)

is ook $SSQ = SSR - [y'y - (SSE + SSQ)]$ nl. $SSR + SSE = y'y$.

Men zou nu kunnen denken dat y'y - (SSE + SSQ) de reductie is van de totale kwadraatsom tengevolge van de modelaanpassing van het gereduceerde model. Dit is evenwel in het algemeen niet waar. Het verschil y'y - (SSE + SSQ) kan zelfs negatief zijn en is dus in het algemeen geen kwadraatsom. De oorzaak hiervan is dat y'y in het gereduceerde model niet altijd de totale kwadraatsom is (zie Searle pag. 117 e.v.).
Echter, in het geval de restrictie luidt $K\beta = 0$ (dus $\gamma = 0$), is y'y wel de totale kwadraatsom bij het gereduceerde model en is y'y - (SSE + SSQ) = = SSR - SSQ =: $SSR_Q$ de reductie hiervan tengevolge van de modelaanpassing. In dat geval is dus

Restkwadraatsom (gereduceerd model) = SSE + SSQ = $SSE_Q$
Reductie in totale kwadraatsom t.g.v. aanpassing (gereduceerd model) = = SSR - SSQ = $SSR_Q$.

Voor het geval $K\beta = 0$ ontstaat de volgende variantie-analyse tabel:

| bron | | | SS | $\nu$ | | verdeling SS |
|---|---|---|---|---|---|---|
| SSR | SSR - SSQ | $SSR_Q$ | $\Vert\hat{y} - \hat{y}_H\Vert^2$ | r | r - q | $\sigma^2\chi^2_\nu(\delta - \delta_Q)$ |
| | SSQ | | $\Vert\hat{y}_H\Vert^2$ | | q | $\sigma^2\chi^2_\nu(\delta_Q)$ |
| SSE | SSE | SSE | $\Vert y - \hat{y}\Vert^2$ | n - r | | $\sigma^2\chi^2_\nu$ |
| SST | SST | SST | $\Vert y\Vert^2$ | n | | $\sigma^2\chi^2_\nu(\delta)$ |

$\hat{y}_H := \hat{y} - \hat{y}_{X_H}$

tabel 7.1.

In deze tabel kunnen twee F-toetsen worden geplaatst:

$H_0$ : $\delta = 0$, $X\beta = 0$, $\beta = 0$ toetst het volle model; $F(R) = MSR/MSE \simeq F^r_{n-r}$ ;

$H_0$ : $\delta_Q = 0$, $K\beta = 0$ toetst de restricties; $F(Q) = MSQ/MSE \simeq F^q_{n-r}$ .

<u>Opmerkingen:</u>

i) Onder $H_0$: $K\beta = 0$ is $SSQ = \Vert V K \underline{b}\Vert^2 = \underline{y}'XX^{-2}K'V^2KX^{-2}X'\underline{y} =: \Vert P_H\underline{y}\Vert^2$, met $\langle H\rangle = \langle XX^{-2}K'\rangle \subset \langle X\rangle$ nl. $P_H$ is symmetrisch, idempotent en $P_H H = H$; (ga na!). $r(P_H) = r(XX^{-2}K') = r(K') = q$.

Definieer $P_{X_H} := P_X - P_H$ dan wordt $\|\hat{y}_{X_H}\|^2 = \|P_{X_H} y\|^2$ en $r(P_{X_H}) = r - q$.

Ga na dat inderdaad $\hat{y}_{X_H} \perp \hat{y} - \hat{y}_{X_H}$

ii) We zien duidelijk de analogie met $P_{X_m} = P_X - P$ en $\hat{y}_m = P_{X_m} y$ etc.

Dit klopt, beschouw nl. de nulhypothese $H_0$ : $\mathcal{E}\bar{\underline{y}} = 0$ oftewel $\mathcal{E}\,\frac{u'}{n}\,\underline{y} =$ $= \frac{u'}{n}\,X\beta = 0$ (kan ook met c iii).

Neem dus $K = u'X/n$ , dus $q = 1$, dan wordt

$$<H> = <XX^{-2}K'> = <XX^{-2}X'u/n> = <P_X u> = <u> \text{ dus } P_H = P \; .$$

En $\hat{y}_m = \hat{y}_{X_H}$ zodat $\hat{y}_{X_H} = \hat{y} - \bar{y}$ en $SSQ = \|\bar{y}\|^2$.

Analoog aan fig. 5.1 kan men de volgende figuur geven:

In het <u>volle model</u> wordt y geprojecteerd op $<X>$ met dim r, in het gereduceerde model op $<X> \cap <H>^{\perp}$ met dim $= r-q$.

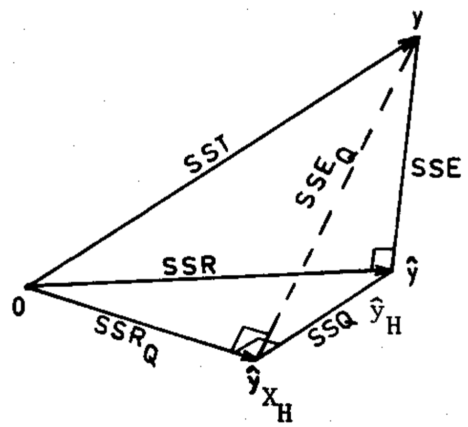


fig. 7.1.

## § 8. Modellen niet van volle rang (d.w.z. $X^2$ singulier)

### a. Inleiding

In de regressie-analyse, besproken in Hoofdstuk I, wordt gewerkt met kwantitatieve verklarende variabelen om het gedrag van een afhankelijke gemeten variabele te verklaren. Vaak komt het evenwel voor dat de verklarende variabelen van kwalitatieve aard zijn. Als we bijvoorbeeld het inkomen van personen meten en proberen te verklaren uit onder meer de genoten opleiding, dan is deze opleiding een kwalitatieve variabele, die in verschillende categorieën, zoals bijvoorbeeld lager onderwijs, lager en middelbaar onderwijs, lager, middelbaar en hoger onderwijs, kan worden verdeeld. Ter onderscheiding van kwantitatieve variabelen zullen we de kwalitatieve variabelen voortaan factoren en de categorieën voortaan niveaus noemen. In de zg. variantie-analyse wordt nu het doel nagestreefd de uitwerking van de niveaus van de factoren op de afhankelijke variabele te analyseren. Analoog aan de regressie-analyse, waar de totale kwadraatsom in een stuk regressiekwadraatsom en een restkwadraatsom werd gesplitst, wordt nu de totale kwadraatsom van de waarnemingen van de afhankelijke variabele gesplitst in stukken die verklaard worden door de verwerkte factoren en een restkwadraatsom. In de klassieke variantie-analyse werkt men daarbij met zgn. orthogonale schema's, d.w.z. bij elke niveaucombinatie van 2 van de aanwezige factoren worden evenveel waarnemingen gedaan. Dit is bij veel problemen, o.a. biologische of economische, vaak niet te realiseren. In die gevallen is de klassieke variantie-analyse niet toepasbaar. Daarom wordt in dit college een andere opzet geïntroduceerd en wordt variantie-analyse als regressie-analyse op dummy variabelen gegeven, die geschikt is voor niet orthogonale schema's en voor orthogonale schema's als speciaal geval van eerstgenoemde. Weliswaar brengt deze methode meer rekenwerk met zich mee, maar dat is door de computer-mogelijkheden geen bezwaar meer. Onder dummy variabelen worden variabelen verstaan die de waarden 0 of 1 kunnen aannemen.

### b. Dummy variabelen

Laten we aan de hand van een voorbeeld het gebruik van dummy variabelen illustreren. De investering van mensen in duurzame gebruiksartikelen wordt gemeten en als verklarende factor wordt de genoten scholing genomen.

Ongetwijfeld zullen er nog andere factoren een verklarende rol kunnen spelen en zal het toeval van invloed zijn op de waarnemingen. Als niveaus voor de factor "scholing" nemen we: lager onderwijs (1), hoger en middelbaar onderwijs (2), en lager, middelbaar en hoger onderwijs (3). Als model nemen we:

8.1. $$\underline{y}_{ij} = \beta_0 + \beta_1 x_{ij,1} + \beta_2 x_{ij,2} + \beta_3 x_{ij,3} + \underline{e}_{ij}$$

waarin

- de $\beta_1$, $\beta_2$, $\beta_3$ de regressie-coëfficiënten zijn van de dummy variabelen
- $\underline{y}_{ij}$ de investering is van de j-de waargenomen persoon op niveau i van de factor "scholing"
- $x_{ij,1}$ de dummy variabel is voor het niveau "lager onderwijs" van de factor "scholing" bij waarneming j, die gelijk is aan 1 als de waargenomen persoon inderdaad alleen lager onderwijs heeft genoten en anders gelijk is aan 0
- analoog $x_{ij,2}$ en $x_{ij,3}$.

Er wordt nu de investering van 6 personen waargenomen waarvan 3 op het laagste, 2 op het middelste en 1 op het hoogste niveau van de factor "scholing". Met het regressiemodel 8.1 vinden we dan:

8.2. $$\begin{aligned}
\underline{y}_{11} &= \beta_0 + \beta_1 && + \underline{e}_{11} \\
\underline{y}_{12} &= \beta_0 + \beta_1 && + \underline{e}_{12} \\
\underline{y}_{13} &= \beta_0 + \beta_1 && + \underline{e}_{13} \\
\underline{y}_{21} &= \beta_0 \quad + \beta_2 && + \underline{e}_{21} \\
\underline{y}_{22} &= \beta_0 \quad + \beta_2 && + \underline{e}_{22} \\
\underline{y}_{31} &= \beta_0 \quad\quad + \beta_3 && + \underline{e}_{31}
\end{aligned}$$

oftewel $\underline{y} = X\beta + \underline{e}$ met $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)'$ en

$$X = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} ; \quad \underline{y} = \begin{pmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{31} \end{pmatrix} ;$$

| | personen j → | | |
|---|---|---|---|
| 1 | $y_{11}$ | $y_{12}$ | $y_{13}$ |
| i ↓ 2 | $y_{21}$ | $y_{22}$ | |
| scholing 3 | $y_{31}$ | | |

Toepassing van de kleinste kwadratenmethode geeft evenals voorheen de normaalvergelijkingen $X'X\underline{b}^0 = X'\underline{y}$. Omdat X niet van volle kolomrang is ($\sum_{2}^{4} x_{*i} = x_{*1}$), is X'X singulier en bestaat $(X'X)^{-1}$ niet. Dit is een voorbeeld dus van een model niet van volle rang.

Een ander voorbeeld is het volgende:

De opbrengst van 6 planten van 3 verschillende rassen behandeld met 2 verschillende soorten kunstmest wordt waargenomen. Er zijn nu twee factoren nl. ras en kunstmest, met resp. 3 en 2 niveaus. We definiëren weer $\underline{y}_{ijk}$ als de opbrengst van de k-de plant van ras i, behandeld met kunstmest j.

Dus $i = 1,2,3$; $j = 1,2$ en $k = 1,\ldots,n_{ij}$. We voeren 5 dummy variabelen in, 3 voor het ras en 2 voor de kunstmest. Omdat we weer naar een regressiemodel op deze dummy variabelen toewillen, hebben we regressiecoëfficiënten nodig. Deze noemen we $\alpha_1$, $\alpha_2$, $\alpha_3$ voor de factor "ras" en $\beta_1$, $\beta_2$ voor de factor "kunstmest". Naast deze hebben we nog een constante term nodig ($\beta_0$ in het vorige voorbeeld) die we $\mu$ zullen noemen. Als model hebben we nu:

8.3.
$$\underline{y}_{ijk} = \mu + \alpha_1 x_{ijk,1} + \alpha_2 x_{ijk,2} + \alpha_3 x_{ijk,3} + \beta_1 x^*_{ijk,1} + \\ + \beta_2 x^*_{ijk,2} + \underline{e}_{ijk} \,.$$

Wanneer de 6 waarnemingen als volgt verdeeld zijn:

| | kunstmest j 1 | 2 |
|---|---|---|
| i ↓ ras 1 | $y_{111}$, $y_{112}$ | $y_{121}$ |
| 2 | $y_{211}$ | $y_{221}$ |
| 3 | $y_{311}$ | |

Dan wordt $\underline{y} = X\beta + \underline{e}$

8.4. $$\begin{pmatrix} y_{111} \\ y_{112} \\ y_{121} \\ y_{211} \\ y_{221} \\ y_{311} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \beta_1 \\ \beta_2 \end{pmatrix} + \underline{e} .$$

Ook nu hebben we een regressiemodel dat helaas niet zo simpel oplosbaar is omdat X niet van volle kolomrang is.

c. <u>Beschrijving van lineaire modellen</u>

De modelvergelijking 8.1 is ontstaan omdat we het lineaire model van de invloed van de factor "scholing" op de investering als regressie op dummy variabelen hebben willen interpreteren. De vergelijkingen 8.2 die hieruit zijn afgeleid kunnen we evenwel met $\mu = \beta_0$ herschrijven tot

8.5. $$\underline{y}_{ij} = \mu + \beta_i + \underline{e}_{ij} .$$

Dit is de gebruikelijke vergelijking van een algemeen lineair model met één factor die zich op een aantal niveaus kan bewegen (index i) en waar bij elk niveau een aantal waarnemingen (index j; $j = 1,\dots,n_i$) zijn gedaan. Hierin heeft $\mu$ de feitelijke betekenis van de algemene gemiddelde investering en representeert $\beta_i$ de specifieke invloed hierop van de beschouwde factor.
De matrix X in 8.2 interpreteren we nu niet als de matrix van dummy variabelen, maar als een "incidentiematrix" waarin 0 en 1 resp. de afwezigheid en aanwezigheid van een bepaald niveau van de factor aangeven.
Vergelijking 8.5 is de vergelijking van een model met <u>enkelvoudige classificatie</u>. Van de $\underline{e}_{ij}$ wordt meestal weer aangenomen dat geldt $\underline{e}_{ij} \sim N(0,\sigma^2)$.
Omgekeerd kan de in de klassieke variantie-analyse gebruikelijke vergelijking 8.5 herleid worden tot 8.1. Beide beschrijven dezelfde situatie. Met deze relatie in het achterhoofd kan de variantie-analyse van het model 8.5 aangepakt worden als een regressie-analyse op dummy variabelen.

Ook het stelsel 8.4 kunnen we anders interpreteren, niet als vergelijkingen in een regressiemodel met dummy variabelen, maar als vergelijkingen van een lineaire variantie-analyse model met 2 factoren:

8.6. $$\underline{y}_{ijk} = \mu + \alpha_i + \beta_j + \underline{e}_{ijk} ,$$

waarbij in ons voorbeeld $i = 1,2,3$; $j = 1,2$; $k = 1,\ldots,n_{ij}$ met $n_{11} = 2$, $n_{12} = 1$, $n_{21} = 1$, $n_{22} = 1$, $n_{31} = 1$, $n_{32} = 0$ (dus $n_{ij}$ = aantal waarnemingen per "cel").

Dit is de vergelijking van een lineair model met tweevoudige classificatie. De $\alpha_i$ en $\beta_j$ geven het effect aan van één niveau van één factor op de respons y. Dit noemt men hoofdeffecten. De $\mu$ is weer een algemeen gemiddelde en de matrix X een incidentiematrix.

In 8.6 zijn de hoofdeffecten additief, d.w.z. het verschil tussen de effecten van kunstmest 1 en 2 op de opbrengst is voor elk ras hetzelfde. Evenzo het verschil tussen de effecten van de rassen 1, 2 en 3 op de opbrengst is voor elke kunstmestsoort hetzelfde. Dit behoeft natuurlijk niet zo te zijn. Het is mogelijk dat verschillende rassen verschillend op kunstmestsoorten reageren. In dat geval spreekt men van interactie tussen de twee factoren en wordt 8.6 uitgebreid met een interactiefactor $(\alpha\beta)_{ij}$:

$$\underline{y}_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \underline{e}_{ijk} .$$

We hebben nu gezien dat de vergelijking voor een algemeen lineair model, zoals opgebouwd in de klassieke variantie-analyse, kan worden geïnterpreteerd als de vergelijking van een regressiemodel op dummy variabelen door de "incidentie-matrix" X als matrix van waarden van dummy variabelen te zien en door de effecten ($\alpha_i$, $\beta_j$ etc.) als regressiecoëfficiënten te beschouwen.

De vergelijking luidt dan $\underline{y} = X\beta + \underline{e}$ en de kleinste kwadratenmethode geeft aanleiding tot de normaalvergelijkingen $X'X\underline{b}^0 = X'\underline{y}$ (zie 5.3).

We schrijven met opzet $\underline{b}^0$ omdat wegens het singulier zijn van $X'X$, er geen éénduidige oplossing van dit stelsel is.

8.7. Veronderstel dat in ons eerste voorbeeld $\underline{y}' = (16,10,19,11,13,27)$ en $\beta' = (\mu,\alpha_1,\alpha_2,\alpha_3)$, dan worden de normaalvergelijkingen:

$$\begin{pmatrix} 6 & 3 & 2 & 1 \\ 3 & 3 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu^0 \\ a_1^0 \\ a_2^0 \\ a_3^0 \end{pmatrix} = \begin{pmatrix} 96 \\ 45 \\ 24 \\ 27 \end{pmatrix} .$$

Merk op dat

$$X'y = \begin{pmatrix} y_{..} \\ y_{1.} \\ y_{2.} \\ y_{3.} \end{pmatrix}$$

en

$$X'X = \begin{pmatrix} n_{..} & n_{1.} & n_{2.} & n_{3.} \\ n_{1.} & n_{1.} & 0 & 0 \\ n_{2.} & 0 & n_{2.} & 0 \\ n_{3.} & 0 & 0 & n_{3.} \end{pmatrix} .$$

Oplossingen hiervan zijn ondermeer: $(\mu^0,a_1^0,a_2^0,a_3^0) = (16,-1,-4,11)$ of $(27,-12,-15,0)$ of $(-2982,2997,2994,3009)$.
Een betekenis kan dus aan een oplossing $b^0$ niet gehecht worden. Als we evenwel de oplossingen $b^0$ nader bekijken, zien we dat $a_1^0 - a_2^0$ voor alle drie dezelfde waarden heeft, evenals bv. $\mu^0 + a_1^0$. De functies van de parameters die invariant zijn voor alle oplossingen $b^0$ heten <u>schatbare functies</u>. Zo'n schatbare functie kan practische betekenis hebben. Bijvoorbeeld $a_1^0 - a_2^0$ is de schatter van het verschil tussen de effecten van 2 niveaus.

d. <u>Normaalvergelijkingen</u>

Evenals in § 5d beschouwen we het model $\underline{y} = X\beta + \underline{e}$.
Echter nu is $X \in \mathcal{M}^r_{n,p}$ met $r < p$, d.w.z. $X'X$ is singulier.

8.8. Het stelsel normaalvergelijkingen is $X'X\underline{b}^0 = X'\underline{y}$.
We schrijven $\underline{b}^0$ i.p.v. $\underline{b}$ omdat er meerdere oplossingen zijn.
(Het stelsel is in elk geval consistent zie 1.50).
Het is dan ook niet juist over $\underline{b}^0$ te spreken als de schatter voor $\beta$ zoals $\underline{b}$ dat was in § 5. Zeker is $\underline{b}^0$ een schatter voor iets, maar niet voor $\beta$. Met $\underline{b}^0$ wordt enkel en alleen een oplossing van de normaalvergelijkingen aangeduid.
Volgens D12 is een oplossing $\underline{b}^0 = (X'X)^- X'y =: GX'y$ met G een gegeneraliseerde inverse van X'X. Passen we dit toe op ons eerste voorbeeld met de normaalvergelijkingen 8.7.
Neem

$$G = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

dan is $\underline{b}^0 = GX'y = (0,15,12,27)'$ en $\hat{y} = Xb^0 = (15,15,15,12,12,27)'$ (vrij evident!). (Controleer $\Sigma\hat{y}_i = \Sigma y_i$, zie 5.14).
Merk op dat juist geldt: $b^0 = (0,\bar{y}_{1.},\bar{y}_{2.},\bar{y}_{3.})'$ en $\hat{y} = (\bar{y}_{1.},\bar{y}_{1.},\bar{y}_{1.},\bar{y}_{2.},\bar{y}_{2.},\bar{y}_{3.})'$

e. Eigenschappen van de oplossingen (vergelijk met § 5g)

i) Verwachtingswaarde

8.9. $\mathcal{E}\underline{b}^0 = GX'\mathcal{E}\underline{y} = GX'X\beta = I^0\beta$ met $I^0 := GX'X$ (zie D1). Dus $\underline{b}^0$ is zuiver voor $I^0\beta$ en niet voor $\beta$ (tenzij $I^0 = I$ dus X volle kolomrang).

ii) Variantie

8.10. var $\underline{b}^0 = GX'\text{var}\,\underline{y}\,XG' = GX'XG'\sigma^2$. In 5.9 was var $\underline{b} = X^{-2}\sigma^2$, dus niet analoog. Dan zou var $\underline{b}^0 = G\sigma^2$ moeten zijn.
Dit kan toch echter bereikt worden door een geschikte keuze van G, nl. volgens D4 kan men X'X transformeren tot een matrix A waarvan de leidende ondermatrix $A_{11}$ regulier is van rang $r = r(X)$.
Daar X'X symmetrisch is, kan dit met dezelfde permutatiematrix $\tilde{I}$:

$$\tilde{I}X'X\tilde{I} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \text{; dan is de gezochte } G = \tilde{I}\begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}\tilde{I}.$$

Ga na dat inderdaad geldt $GX'XG' = G$ en dus var $\underline{b}^0 = G\sigma^2$.
Men kan deze G direct vinden, nl. de Penrose inverse $(X'X)^+$ voldoet. Daarvoor geldt nl. $B^+BB^+ = B^+$ en $B^+$ is symmetrisch als B het is (zie Da).

iii) Schatter voor $\mathcal{E}\underline{y} = X\beta$

8.11. Zij $\hat{\underline{y}}$ de schatter voor $\mathcal{E}\underline{y}$, dan is $\hat{\underline{y}} = X\underline{b}^0 = XGX'\underline{y} =: P_X\underline{y}$ (analoog 5.11). Volgens D10 is $P_X$ onafhankelijk G dus $\hat{y}$ onafhankelijk $\underline{b}^0$ en dus kan $\hat{\underline{y}}$ gezien worden als de schatter voor $\mathcal{E}\underline{y}$.
Deze is bovendien zuiver daar $\mathcal{E}\hat{\underline{y}} = P_X X\beta = X\beta$.

f. Restkwadraatsom

Analoog aan 5.13 is $SSE := \| y - \hat{y} \|^2 = \| P_X^\perp y \|^2$, dus eveneens onafhankelijk $\underline{b}^0$.
Analoog geldt $SSE = \underline{y}'\underline{y} - \underline{b}^{0'}X'\underline{y}$.
Eveneens als 5.15 is $\underline{s}^2 = SSE/(n - r)$ een zuivere schatter voor $\sigma^2$ en onafhankelijk $\underline{b}^0$. De bewijzen behoeven niet opnieuw gegeven te worden!
In het model van volle rang was $P_X = X(X'X)^{-1}X'$ en nu $P_X = XGX'$.

g. Splitsen van de totale kwadraatsom

Dit gebeurt geheel analoog als bij het model van volle rang (zie § 5ℓ). Het enige verschil is dat het hier geen zin heeft te werken met x-variabelen die gecorrigeerd zijn t.o.v. het gemiddelde.
Zo wordt bv. $SSR_m = \| P_{X_m} y \|^2 = y'(XGX' - uu'/n)y$ etc.
Tabel 6.1 en 6.2 blijven ongewijzigd. Voor de verdelingen der SS zie 8.14.
Passen we dit alles toe op ons eerste voorbeeld 8.7 met G als in 8.8

$$G = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \text{ dan wordt } I^0 = GX'X = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix};\ b^0 = (0,15,12,27)',$$

$n = 6$, $\bar{y}_. = 16$; $r(X) = 3 = r$.

$SSR = b^{0'}X'y \quad = 1692$ ;

$SST = y'y \quad = 1736 \qquad SST_m = SST - SSM = 200$

$SSM = n\bar{y}_.^2 \quad = 1536 \qquad SSR_m = SSR - SSM = 156$ .

$SSE = SST - SSR = 44$

Zo ontstaat:

| bron | SS | $\nu$ | MS |
|---|---|---|---|
| SSR | 1692 | 3 | 564 |
| SSE | 44 | 3 | 14.7 |
| SST | 1736 | 6 | |

of

| bron | SS | $\nu$ | MS |
|---|---|---|---|
| $SSR_m$ | 156 | 2 | 78 |
| SSE | 44 | 3 | 14.7 |
| $SST_m$ | 200 | 5 | |

h. Verdelingen in het niet volle rang-model

Ons model is weer $\underline{y} = X\beta + \underline{e}$ met $\underline{e} \sim N_n(0,\sigma^2 I)$ en $X \in \mathfrak{M}^r_{n,p}$ met $r < p \leq n$.

We weten $\underline{b}^0 = GX'\underline{y}$; Nu is $(GX')_{p\times n}$; $r(GX') \leq r(X') = r < p$ dus $GX'$ is niet van volle rijrang en theorema 3.3 niet van toepassing.
Wel volgt uit dit theorema, dat

8.12. $$\varphi_{\underline{b}^0}(t) = \exp[it'I^0\beta - \tfrac{1}{2}t'GX^{\underline{2}}G't\sigma^2] \text{ (met 8.9 en 8.10)} .$$

Omdat $GX'$ niet van volle rijrang is, is $GX'XG'$ singulier nl. $r(GX'XG') = r(XG')^{\underline{2}} = r(XG') = r(GX')$. Overigens lijkt 8.12 op de karakteristieke functie van een multinormale verdeling.
We noemen de verdeling van $\underline{b}^0$ dan ook singulier normaal, notatie $N^0_p$, dus

8.13. $$\underline{b}^0 \sim N^0_p(I^0\beta, I^0G'\sigma^2) .$$

Opmerking: Nemen we G symmetrisch, volgens D9 altijd mogelijk, dan ontstaat

$$\underline{b}^0 \sim N^0_p(I^0\beta, I^0G\sigma^2) \text{ analoog aan } \underline{b} \sim N_p(\beta,(X'X)^{-1}\sigma^2) .$$

Deze gaan in elkaar over als $I^0 = I$ oftewel $r = p$.

8.14. De verdelingen der SS in de variantie-analyse tabel veranderen niet, daar deze alleen afhangen van de projecties en hun verwachtingen en deze zijn dezelfde gebleven als in het model van volle rang (behalve dat we b door $b^0$ en $(X'X)^{-1}$ door G vervangen). Ook is de gehele analyse onafhankelijk $b^0$ (en G) daar de projecties dat zijn.

k. Het toetsen van hypothesen (vergelijk § 6d)

In het volle rang-model werd met F(R) getoetst: $X\beta = 0$ dus $\beta = 0$.
In het niet volle rang-model toetst F(R): $X\beta = 0$ (dus niet $\beta = 0$).
Dit volgt ook uit het feit dat de elementen van $\beta$ niet schatbaar zijn, zoals we nog zullen zien. Er is geen sprake van de schatter van $\beta$. Wèl getoetst kan worden met F(R), $H_0 : X\beta = 0$. Immers $X\beta$ is een schatbare functie nl. $\hat{\underline{y}} = X\underline{b}^0$ is een schatter voor $X\beta$ die, zoals we zagen, invariant is m.b.t. $\underline{b}^0$.
Analoog wordt met $F(R_m)$ getoetst $H_0 : X_m\beta_m = 0$ (en niet $\beta_m = 0$).

Passen we dit weer toe op ons voorbeeld, zie § 8g.

$F(M) = MSM/MSE = 104$ is duidelijk significant zodat $H_0 : \mathcal{E}\bar{\underline{y}} = 0$ wordt verworpen! Echter $F(R_m) = MSR_m/MSE = 5.32 < F_3^2(\alpha) = 5.46$ voor $\alpha = 0{,}10$ dus nog geen significant resultaat.
$H_0 : X_m\beta_m = 0$ kan niet verworpen worden.
Bekijken we deze hypothese nog eens nader:

$$X_m = X_1 - \bar{X}_1 = \begin{pmatrix} \frac{1}{2} & -\frac{1}{3} & -\frac{1}{6} \\ \frac{1}{2} & -\frac{1}{3} & -\frac{1}{6} \\ \frac{1}{2} & -\frac{1}{3} & -\frac{1}{6} \\ -\frac{1}{2} & \frac{2}{3} & -\frac{1}{6} \\ -\frac{1}{2} & \frac{2}{3} & -\frac{1}{6} \\ -\frac{1}{2} & -\frac{1}{3} & \frac{5}{6} \end{pmatrix} ; \quad \beta_m = (\alpha_1, \alpha_2, \alpha_3)'.$$

$$r(X_m) = 2$$

Het stelsel vergelijkingen $X_m\beta_m = 0$ reduceren we tot

$$\begin{pmatrix} \frac{1}{2} & -\frac{1}{3} & -\frac{1}{6} \\ -\frac{1}{2} & \frac{2}{3} & -\frac{1}{6} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

waaruit volgt: $\alpha = \lambda(1,1,1)$ oftewel $H_0 : \alpha_1 = \alpha_2 = \alpha_3$.
Dus de nulhypothese $H_0 : X_m\beta_m = 0$ houdt in dit geval in: $\alpha_1 = \alpha_2 = \alpha_3$ oftewel: geen verschil in niveaus!

## § 9. Schatbare functies

a. Het idee van schatbare functie is al ter sprake gekomen. In principe is het een lineaire functie van de parameters waarvoor een schatter bestaat die invariant is met betrekking tot de gekozen oplossing $\underline{b}^0$ van de normaalvergelijkingen.

9.1. Definitie: Een lineaire functie $q'\beta$ van de parameters $\beta$ heet schatbaar als er een vector t bestaat zodat $q'\beta = t'\mathcal{E}\underline{y}$ (oftewel $\mathcal{E}(t'\underline{y}) = q'\beta$).

Hierbij behoeft t beslist niet eenduidig bepaald te zijn.
In ons voorbeeld is $\mu + \alpha_1 = \mathcal{E}\underline{y}_{11}$ een schatbare functie, nl.

$$\mu + \alpha_1 = (1,1,0,0)\begin{pmatrix}\mu\\ \alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix} = q'\beta \text{ en } \mathcal{E}\underline{y}_{11} = \delta'_{1*}\mathcal{E}\underline{y} = t'\mathcal{E}\underline{y}.$$

Maar ook $\mu + \alpha_2 = \mathcal{E}\underline{y}_{21}$ en dus ook $\alpha_1 - \alpha_2 = \mathcal{E}(\underline{y}_{11} - \underline{y}_{21})$ zijn schatbare functies. $\mu + \alpha_1$ is ook $= \mathcal{E}\underline{y}_{12} = \mathcal{E}\underline{y}_{13}$ waaruit de niet-eenduidigheid van t blijkt.

Eigenschappen.

9.2. De verwachting van elke waarneming $\eta_i$, dus $\mathcal{E}\underline{y}_i$, is een schatbare functie, nl. neem $t = \delta_{i*}$ dan is $\mathcal{E}(t'\underline{y}) = \mathcal{E}\underline{y}_i = (X\beta)_i = x'_{i*}\beta = q'\beta$.

9.3. Lineaire combinaties van schatbare functies zijn weer schatbare functies, nl. stel $q_1'\beta$ en $q_2'\beta$ zijn schatbare functies, d.w.z. $\exists t_1$, $\exists t_2$ zodat $q_1'\beta = \mathcal{E}t_1'\underline{y}$ en $q_2'\beta = \mathcal{E}t_2'\underline{y}$. Dus geldt ook

$$c_1q_1'\beta + c_2q_2'\beta = c_1\mathcal{E}t_1'\underline{y} + c_2\mathcal{E}t_2'\underline{y} = (c_1t_1' + c_2t_2')\mathcal{E}\underline{y} =: t'\mathcal{E}\underline{y}\ .$$

Analoog voor meerdere schatbare functies.

9.4. <u>Stelling</u>: $q'\beta$ schatbaar $\Longleftrightarrow$ $q \in \langle X' \rangle$ (oftewel $q' = t'X$).

<u>Bewijs</u>:

$\Longrightarrow$: $q'\beta = t'\mathcal{E}\underline{y} = t'X\beta$ dus $q' = t'X$ aangezien de schatbaarheid van een functie niet van de parameterwaarden $\beta$ afhangt en dus geldt voor alle $\beta$.
$\Longleftarrow$: Stel $q' = t'X$, dus $q'\beta = t'X\beta = t'\mathcal{E}\underline{y}$ oftewel $q'\beta$ schatbaar.

Daar X grote afmetingen kan hebben, waardoor het vinden van een t bij een gegeven q lastig kan zijn, werken we liever met $I^0 = GX'X$ daar deze een eenvoudigere vorm heeft. Zo volgt:

9.5. <u>Stelling</u>: $q'\beta$ schatbaar $\Longleftrightarrow$ $q'I^0 = q'$.

<u>Bewijs</u>: Dit volgt direct uit 9.4, nl. d.e.s.d. geldt:

$$q' = t'X = t'XI^0 = q'I^0 \text{ daar } XI^0 = X \text{ (D10)} .$$

<u>Opmerking</u>: Is X van volle rang dan is $I^0 = I$, dus is $q'\beta$ schatbaar voor elke q.

9.6. <u>Stelling</u>: Is $q'\beta$ schatbaar, dan is $q'\underline{b}^0$ onafhankelijk $\underline{b}^0$ (dus onafhankelijk G).

<u>Bewijs</u>: Dus $q' = t'X$; $q'\underline{b}^0 = t'X\underline{b}^0 = t'P_X\underline{y}$, onafhankelijk G, dus ook onafhankelijk $\underline{b}^0$. Dit is het belangrijkste aspect van schatbare functies, het feit namelijk dat diens schatter onafhankelijk is van de gekozen oplossing! Gevolg is natuurlijk dat $\text{var}(q'\underline{b}^0)$ ook onafhankelijk $\underline{b}^0$ en G is!

<u>Opmerking</u>: Het schatbaar zijn van een functie $q'\beta$ impliceert dat er een t is zodat $q' = t'X$ oftewel $X't = q$. Nu, als dit stelsel dus consistent is, is er een $(n - r)$ dim. variëteit van oplossingen, dus t is zeker niet eenduidig bepaald!
Daar $q'\underline{b}^0 = t'X\underline{b}^0$ kan de vraag rijzen of deze schatter afhankelijk is van t. Het antwoord is natuurlijk "neen" omdat, wèlke t we ook bij gegeven vaste q nemen, steeds geldt $q' = t'X$ en dus $t'X\underline{b}^0 = q'\underline{b}^0$.

9.7. <u>Stelling</u>: $q'\underline{b}^0$ is BLUE voor $q'\beta$ (schatbaar) (zie § 5f).

<u>Bewijs</u>:

i) $q'\underline{b}^0 = q'GX'\underline{y}$ dus een lin. functie der waarnemingen (lineair estimator)

ii) $\mathcal{E}q'\underline{b}^0 = q'\mathcal{E}\underline{b}^0 = q'I^0\beta = q'\beta$ (met 9.5) (unbiased)

iii) $\mathrm{var}(q'\underline{b}^0) = \mathrm{var}(t'X\underline{b}^0) = \mathrm{var}(t'P_X\underline{y}) = t'P_Xt\sigma^2 = t'XGX't\sigma^2 = q'Gq\sigma^2$.
Onderstel dat $k'\underline{y}$ ook een lineaire zuivere schatter is voor $q'\beta$.
Dus moet $\mathcal{E}k'\underline{y} = k'\mathcal{E}\underline{y} = k'X\beta = q'\beta$ oftewel $q' = k'X$.
$\mathrm{cov}(q'\underline{b}^0,k'\underline{y}) = \mathrm{cov}(t'P_X\underline{y},k'\underline{y}) = t'P_Xk\sigma^2 = t'XGX'k\sigma^2 = q'Gq\sigma^2$.
Dus: $0 < \mathrm{var}(q'b^0 - k'\underline{y}) = \mathrm{var}(q'\underline{b}^0) + \mathrm{var}(k'\underline{y}) - 2\,\mathrm{cov}(q'\underline{b}^0,k'\underline{y}) =$
$= \mathrm{var}(k'\underline{y}) - \mathrm{var}(q'\underline{b}^0)$ oftewel $\mathrm{var}(k'\underline{y}) > \mathrm{var}(q'\underline{b}^0)$ (best). q.e.d.

Zijn de componenten $q_{i*}'\underline{b}^0$ van $Q\underline{b}^0$ de BLUE's van de schatbare functies $q_{i*}'\beta$ (componenten van $Q\beta$) dan is

$$\mathrm{VAR}(Q\underline{b}^0) = \mathrm{VAR}(TX\underline{b}^0) = \mathrm{VAR}(TP_X\underline{y}) = TP_XT'\sigma^2 = TXGX'T'\sigma^2 = QGQ'\sigma^2 .$$

<u>Betrouwbaarheidsinterval</u>

Aangezien $q'\underline{b}^0 = q'GX'\underline{y}$ en $g'GX'$ volle rijrang heeft is (3.3)
$q'\underline{b}^0 \sim N(q'\beta,q'Gq\sigma^2)$. Analoog aan 6.8 wordt het betrouwbaarheidsinterval:

$$q'\underline{b}^0 \pm t_{n-r}(\tfrac{1}{2}\alpha)\underline{s}\sqrt{q'Gq} .$$

b. <u>Welke functies zijn schatbaar?</u>

Schatbare functies zijn lineaire functies van de vorm $q'\beta$ met $q \in \langle X'\rangle$.
Daar $\dim\langle X'\rangle = r$, zijn er hoogstens r lineair onafhankelijke vectoren $q_1,\ldots,q_r$ en dus ook hoogstens r lineair onafhankelijke schatbare functies. Alle andere vectoren $q'$ zijn lineaire combinaties van deze r onafhankelijke vectoren en dus <u>alle</u> overige schatbare functies zijn lineaire combinaties van de bij $q_1,\ldots,q_r$ behorende schatbare functies.
Deze verzameling van schatbare functies wordt ook verkregen wanneer we als algemene vorm nemen $m'X\beta$, m willekeurig, of $v'X'X\beta$, v willekeurig of ook $w'\mathcal{E}\underline{b}^0 = w'GX'X\beta$, w willekeurig. Deze lineaire functies zijn alle schatbaar volgens de definitie (neem $m'X$, $v'X'X$, $w'GX'X$ resp. $q'$), nl.

i) Dat $m'X\beta$, m willekeurig alle schatbare functies genereert evenals $q'\beta$ met $q' = t'X$, t willekeurig, is triviaal.

ii) Daar $\langle X'X\rangle = \langle X'\rangle$ zijn er ook r onafhankelijke vectoren van het type $v_1'X'X,\ldots,v_r'X'X$. Daardoor worden met $v'X'X\beta$ evenals met $q'\beta$ alle schatbare functies gevonden.

iii) Daar $\langle X'XG'\rangle = \langle X'\rangle$ (nl. $\langle X'XG'\rangle \subset \langle X'\rangle$ en $r(X'XG') = r(GX'X) = r$ (zie D1)) zijn er ook r onafhankelijke vectoren van het type $w_1'GX'X,\ldots,w_r'GX'X$. Daardoor worden met $w'GX'X\beta = w'\mathcal{E}\underline{b}^0$ evenals met $q'\beta$ alle schatbare functies gevonden.

Tabel 9.1.

| Schatbare functie | BLUE | variantie van deze schatter |
|---|---|---|
| $q'\beta$ $(q' = t'X)$ | $q'\underline{b}^0$ | $q'Gq\sigma^2$ |
| $m'X\beta$ | $m'X\underline{b}^0$ | $m'P_X m\sigma^2$ |
| $v'X'X\beta$ | $v'X'X\underline{b}^0 = v'X'\underline{y}$ | $v'X'Xv\sigma^2$ |
| $w'\mathcal{E}\underline{b}^0 = w'GX'X\beta$ | $w'\underline{b}^0$ | $w'I^0G'w\sigma^2$ |

Opmerkingen:

i) De BLUE voor $v'X'X\beta$ is $v'X'X\underline{b}^0 = v'X'XGX'\underline{y} = v'X'\underline{y}$.

ii) De BLUE voor $w'\mathcal{E}\underline{b}^0$ is $w'GX'X\underline{b}^0 = w'GX'\underline{y} = w'\underline{b}^0$.

iii) Nemen we in $w'\mathcal{E}\underline{b}^0$, $w = \delta_{i*}$ dan geldt dat $\delta_{i*}'\underline{b}^0 = \underline{b}_i^0$, de BLUE is voor de schatbare functie $(\mathcal{E}\underline{b}^0)_i = (I^0\beta)_i$.
We noemen dan per definitie de vector $\underline{b}^0$ BLUE voor $I^0\beta$.

iv) Het geval $v'X'X\beta$ geeft op een eenvoudigere manier dan $m'X\beta$ alle schatbare functies omdat v slechts p en m echter n elementen heeft.
Het geval $w'GX'X\beta$ is evenwel weer eenvoudiger dan $v'X'X\beta$ vanwege de eenvoudige schatter $w'\underline{b}^0$ en de vaak eenvoudige vorm van $I^0 := GX'X$.

c. Voorbeeld

Passen we dit laatste toe op ons voorbeeld uit § 8g, dan vinden we voor de schatbare functies de volgende uitdrukkingen:

i) $m'X\beta = \mu \sum_{1}^{6} m_i + \alpha_1 \sum_{1}^{3} m_i + \alpha_2 \sum_{4}^{5} m_i + \alpha_3 m_6.$

De BLUE voor $m'X\beta$ is $m'X\, b^0$ wordt

$$= 15 \sum_{1}^{3} m_i + 12 \sum_{4}^{5} m_i + 27 m_6 \ .$$

Voor willekeurige $m_1,\ldots,m_6$ hebben we hier een schatbare functie met bijbehorende schatting.

ii) $v'X'X\beta = \mu(6v_1 + 3v_2 + 2v_3 + v_4) + \alpha_1(3v_1 + 3v_2) + \alpha_2(2v_1 + 2v_3) + $ $+ \alpha_3(v_1 + v_4).$

De BLUE wordt $v'X'Xb^0 = 15(3v_1 + 3v_2) + 12(2v_1 + 2v_3) + 27(v_1 + v_4).$

Voor willekeurige $v_1,\ldots,v_4$ hebben we weer een schatbare functie met bijbehorende schatting.

iii) $w'GX'X\beta = w'I^0\beta = \mu(w_2 + w_3 + w_4) + \alpha_1 w_2 + \alpha_2 w_3 + \alpha_3 w_4.$

De BLUE wordt $w'GX'Xb^0 = w'b^0 = 15w_2 + 12w_3 + 27w_4.$

Voor willekeurige $w_2$, $w_3$ en $w_4$ hebben we een schatbare functie met bijbehorende schatting.

Daar met alle drie alle schatbare functies gevonden kunnen worden, blijkt hieruit duidelijk dat het derde geval het eenvoudigst is.

## § 10. De algemene lineaire hypothese in het model niet van volle rang

### a. "Toetsbare" hypothesen

Onder een "toetsbare" hypothese wordt een hypothese verstaan die bestaat uit schatbare functies. We zullen later zien dat een hypothese die niet uit schatbare functies is samengesteld, niet te toetsen is en beschouwen daarom nu alleen toetsbare hypothesen.

10.1. Definitie: Een hypothese $H : K\beta = \gamma$ is toetsbaar als $K\beta$ uit q schatbare functies $k'_{i*}\beta$ bestaat.

10.2. Opmerkingen:

i) Omdat $k'_{i*}\beta$ schatbaar is geldt $k'_{i*} = t'_iX$ en dus $K = TX = TXI^0 = KI^0$, waarin T een matrix is van de orde $q \times n$.

ii) Omdat bij het toetsen van een hypothese steeds uitgegaan wordt van lineair onafhankelijke componenten, moeten de schatbare functies $k'_{i*}\beta$ $(i = 1,\ldots,q)$ lineair onafhankelijk zijn oftewel heeft K volle rijrang q. Omdat er bovendien hoogstens r (rang van X) lineair onafhankelijke schatbare functies zijn, geldt $q \leq r < p$.

iii) De vector $K\beta$ bevat q schatbare functies. Hun BLUE's worden gegeven door de corresponderende elementen van de vector $K\underline{b}^0$.
Er geldt dus:

$K\underline{b}^0$ bevat de BLUE's voor $K\beta$ ;

10.3. $$\mathcal{E}K\underline{b}^0 = K\mathcal{E}\underline{b}^0 = KI^0\beta = K\beta \text{ (zie 10.2)}$$

$$\mathrm{var}(K\underline{b}^0) = K \,\mathrm{var}\, \underline{b}^0K' = KI^0G'K'\sigma^2 = KG'K'\sigma^2 = TXG'X'T'\sigma^2 =$$
$$= KGK'\sigma^2 =: V^{-2}\sigma^2 \ .$$

KGK' is regulier nl: KGK' is van de orde $q \times q$; K is van volle rijrang (zie ii).

10.4. KGK' is onafhankelijk G, dus neem $G > 0$ (bestaat volgens D11) en dus is volgens 1.66 KGK' regulier.

b. Het toetsen van "toetsbare" hypothesen

De toets voor het toetsen van de "toetsbare" hypothese H : $K\beta = \gamma$ met $K \in \mathcal{M}^q_{q,p}$ wordt op dezelfde wijze ontwikkeld als in het geval van het model van volle rang (§ 7). Aangenomen wordt weer $\underline{e} \sim N_n(0,\sigma^2 I)$ en $\underline{y} \sim N_n(X\beta,\sigma^2 I)$.
We hebben een toetsingsgrootheid nodig en beschouwen daartoe eerst $K\underline{b}^0$.
Evenals in § 7a geldt:

10.5. $$K\underline{b}^0 - \gamma \sim N_q(K\beta - \gamma,\ V^{-2}\sigma^2) \quad \text{met } V^{-2} := KGK' ,$$

nl. $K\underline{b}^0 = KGX'\underline{y}$; Nu is $r(KGX') = r(KGX'XG'K') = r(KGX'XG'X'T') = r(KGX'T') = r(KGK') = q$ (10.4). Dus theorema 3.3 is toepasbaar en met 10.3 is dus 10.5 bewezen.
Analoog definiëren we SSQ := $\| V(K\underline{b}^0 - \gamma)\|^2$.
Verder kan verwezen worden naar § 7a,b daar niets verandert.
Alleen natuurlijk b door $b^0$ en $(X'X)^{-1}$ door G vervangen.

c. Speciale lineaire hypothesen

De speciale gevallen van $K\beta = \gamma$ die in het model van volle rang zijn onderzocht, o.a. $\beta = 0$, kunnen we nu niet bekijken omdat de $\beta_i$'s geen schatbare functies zijn. Immers $\beta = K\beta$ met $K = I$ en $II^0 \neq I$ (10.2) dus zijn de $\beta_i$'s geen schatbare functies.
Een wèl veel voorkomend geval van een speciale hypothese is H : $K\beta = 0$ met $KI^0 = K$ dus $k_{i*}'\beta$ schatbare functies $i = 1,\dots,q$.
Onder H : $K\beta = 0$ wordt SSQ = $\| VK\underline{b}^0\|^2$ en

$$\underline{b}^0_{X_H} = \underline{b}^0 - GK'V^2K\underline{b}^0 \quad \text{(analoog aan 7.5)} .$$

Zo ontstaat dezelfde variantie-analyse tabel als tabel 7.1.
Deze kan men ook weer corrigeren voor het gemiddelde.
De volgende toetsen kan men bekijken:

F(R) = MSR/MSE toetst $H_0$ : $X\beta = 0$ (het volle model)
F(Q) = MSQ/MSE toetst $H_0$ : $K\beta = 0$ (de restricties)
$MSR_Q$/MSE toetst het gereduceerde model $X\beta = 0$ onder de restricties $K\beta = 0$
$MSR_m$/MSE toetst het volle model $X_m\beta_m = 0$, na aanpasssing van het gemiddelde.

<u>Opmerking</u>: $X_m\beta_m$ bestaat uit schatbare functies nl.

$$X_m\beta_m = (0 \quad X_1 - \bar{X}_1)\beta = (X - \bar{X})\beta = P^{\perp}X\beta =: K\beta \text{ met } P^{\perp}XI^0 = P^{\perp}X\ .$$

Echter $P^{\perp}X = K$ is niet van volle rijrang, d.w.z. de schatbare functies zijn afhankelijk. Een zelfde opmerking geldt voor $X\beta$.

d. <u>Niet-toetsbare hypothesen</u>

We zagen onder 10.2 dat $H : K\beta = \gamma$ toetsbaar is als $KI^0 = K$.

10.6. <u>Definitie</u>: $H : K\beta = \gamma$ is <u>niet</u> toetsbaar als $\langle K'\rangle \cap \langle X'\rangle = \{0\}$ oftewel: iedere vector in de rijenruimte van K is lineair onafhankelijk van de vectoren in de rijenruimte van X. (Opm. dit is niet de juiste ontkenning van het begrip toetsbaar) Dus $K\beta$ bestaat uit q niet schatbare functies.

We gaan nu aantonen dat indien $H : K\beta = \gamma$ gëén toetsbare hypothese is, $SSE_Q = SSE$ oftewel $SSQ = 0$ zodat de toetsingsgrootheid vervalt.
We trachten $\|y - X\beta\|^2$ te minimaliseren onder de voorwaarde $K\beta = \gamma$.
Analoog aan § 7b vinden we de vergelijkingen $X'Xb^0_{X_H} + K'\theta = X'y$ ; $Kb^0_{X_H} = \gamma$.
De bewering is: er is voor dit stelsel vergelijkingen een oplossing $b^0_{X_H}$ die tevens oplossing is van de oorspronkelijke normaalvergelijkingen $X'Xb^0 = X'y$. Dit zou inhouden dat het toevoegen van de restricties geen invloed uitoefent op het te bereiken minimum van de restkwadratensom oftewel $SSQ = 0$.

Om de bewering te bewijzen, beschouw het stelsel vergelijkingen:

10.7. $$K(I^0 - I)z = \gamma - KGX'y\ .$$

Dit stelsel is <u>consistent</u> o.a. als $K(I^0 - I)$ van volle rijrang is, dus q. Dit is zo nl.

i) $r(I^0 - I) = r(I^0 - I)' = p - r$ (zie D2).

ii) $N(I^0 - I)' = \langle X'\rangle$ nl. $\dim N(I^0 - I)' = r = \dim\langle X'\rangle$ en bovendien $\langle X'\rangle \subset N(I^0 - I)'$ daar $(I^0 - I)'X' = (X'XG' - I)X' = 0$.

iii) $r[K(I^0 - I)] = r[(I^0 - I)'K'] = q$ nl. los op $(I^0 - I)'K'x = 0,\ x \neq 0 \rightarrow$ $\rightarrow K'x \neq 0$ daar $K'$ van volle kolomrang is. Eveneens geldt $K'x \notin N(I^0 - I)' = \langle X'\rangle$ daar gegeven is dat $\langle K'\rangle \cap \langle X'\rangle = \{0\}$, oftewel alleen $x = 0$ voldoet dus rang = q.

iv) Alle oplossingen van $X'Xb = X'y$ zijn te schrijven als (zie D13)

10.8. $$b^0 = GX'y + (I^0 - I)z \,.$$

Neem voor $z$ een oplossing $z_1$ van het stelsel 10.7 en noem dan de oplossing van 10.8 $b^0_{X_H} = GX'y + (I^0 - I)z_1$.
Nu blijkt $Kb^0_{X_H} = KGX'y + K(I^0 - I)z_1 = \gamma$ (uit 10.7).

Conclusie: de $Kb^0 = \gamma$ geven geen extra restricties. Er blijft een oplossing voor de oorspronkelijke normaalvergelijkingen die tevens aan de restricties voldoet, en dus dat $SSE_Q = SSE$ oftewel $SSQ = 0$.

Opmerking: Als we niet van te voren hebben nagegaan of de niet-toetsbare hypothese $K\beta = \gamma$ toetsbaar is door te onderzoeken of het aan het criterium $KI^0 = K$ voldoet en als $KGK'$ toch regulier is, wat mogelijk is, dan krijgen we toch een $SSQ \neq 0$. Evenwel is na te gaan dat we dan in feite niet $H : K\beta = \gamma$ toetsen maar de hypothese $H : KI^0\beta = \gamma$.

e. Voorbeeld

We gaan terug naar ons voorbeeld.

i) $H_0 : \alpha_1 = \alpha_2 + 5$.
We hebben al gezien dat $\alpha_1 - \alpha_2$ een schatbare functie is.
Toepassing van het criterium 9.5 laat dit nogmaals zien:

$$K = (0,1,-1,0) \;;\; I^0 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \;;\; KI^0 = K \,.$$

$Kb^0 - \gamma = -2$ (nl. $\gamma = 5$; $b^0 = (0,15,12,27)'$ zie 8.8).
$KGK' = 5/6$ zodat $SSQ = -2(5/6)^{-1}(-2) = 24/5$.
De toetsingsgrootheid wordt:
$F(Q) = MSQ/MSE = \dfrac{24/5}{44/3} = 0.32$, niet significant, $H_0$ niet verwerpen.

ii) $H_0 : \mu + \alpha_1 = 10$ oftewel $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 10 \\ 10 \end{pmatrix}$ oftewel $K\beta = \gamma$.
$\mu + \alpha_2 = 10$

Dat dit een toetsbare hypothese is volgt weer uit het feit dat $KI^0 = K$. Nu wordt

$$Kb^0 - \gamma = \begin{pmatrix} 15 \\ 12 \end{pmatrix} - \begin{pmatrix} 10 \\ 10 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \end{pmatrix} \text{ en } KGK' = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

zodat

$$SSQ = (5,2)\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 5 \\ 2 \end{pmatrix} = 83 \text{ en}$$

$$F(Q) = \frac{83/2}{44/3} = 2.83,$$

een trekking uit $F_3^2$; niet significant, $H_0$ niet verworpen.

## § 11. Beperkingen op het model of op de oplossingen

### a. Beperkingen op het model

Een lineair model kan soms beperkingen inhouden wat betreft de parametervector $\beta$. Deze beperkingen verschillen wezenlijk van de "gebruikelijke" beperkingen (usual constraints) op de oplossingen van de normaalvergelijkingen die alleen ingevoerd worden om één oplossing te verkrijgen. Deze laatste worden onder b besproken.
Tot nog toe is wat slordig gepraat over de vector van parameters $\beta$. Er is geen formele definitie gegeven. In ons voorbeeld is $\mu$ beschreven als het "algemeen gemiddelde" en $\alpha_1$, $\alpha_2$, $\alpha_3$ als "de effecten van de scholing". Soms evenwel volgen uit scherpere definities van de parameters, die inherent zijn aan het model, bepaalde relaties tussen of beperkingen op deze parameters. Bijvoorbeeld kan $\mu$ gedefinieerd worden als het gemiddelde van de verwachtingen op de 3 schoolniveaus:

$$\underline{y}_{ij} = \mu + \alpha_i + \underline{\varepsilon}_{ij}, \quad i = 1,2,3 \; ; \; j = 1,\ldots,n_i$$

$$\mu := \frac{1}{3}\mathcal{E}(\underline{y}_{1j} + \underline{y}_{2k} + \underline{y}_{3\ell}) .$$

Dan volgt daaruit $\mu = \frac{1}{3}(\mu + \alpha_1 + \mu + \alpha_2 + \mu + \alpha_3)$ oftewel $\alpha_1 + \alpha_2 + \alpha_3 = 0$. Deze relatie wordt dan ook niet als een hypothese, maar als een feit bij de verdere analyse aanvaard. Dergelijke relaties heten beperkingen op het model. Modellen zonder dergelijke beperkingen noemt men niet-beperkte modellen, in het andere geval spreekt men van beperkte modellen.
Hoe verloopt nu in een beperkt model het bepalen van een oplossing voor de normaalvergelijkingen $X'X\underline{b} = X'\underline{y}$?
Beschouwen we in het algemeen het volgende stelsel beperkingen:

11.1. $\quad B\beta = \delta$ met $B \in \mathfrak{M}^{q}_{q,p}$ dus van volle rijrang.

De beperkingen $B\beta = \delta$ worden als deel van het model opgevat zodat

$$\underline{y} = X\beta + \underline{e}$$
$$B\beta = \delta$$

het beperkte model vormen.

Het minimaliseren van $\|y - X\beta\|^2$ met de kleinste kwadratenmethode onder deze restricties geeft, met $2\theta$ als de vector van de Lagrange multiplicatoren, weer aanleiding tot de vergelijkingen

11.2. $$X'Xb_r^0 + B'\theta = X'y \qquad \text{(vergelijk met 7.2 en 7.3)}$$

$$Bb_r^0 = \delta \ .$$

Hierin is $b_r^0$ de oplossing van de normaalvergelijkingen van het beperkte model.

We beschouwen nu twee gevallen:

i) De elementen van $B\beta$ zijn schatbaar.

Dan is dit geval volkomen analoog aan dat van de toetsbare hypothese $K\beta = \gamma$ en vinden we analoog aan 7.5

$$\underline{b}_r^0 = \underline{b}^0 - GB'V^2(B\underline{b}^0 - \delta) \text{ met } V^2 := (BGB')^{-1} \ .$$

Eveneens $SSE_r = SSE + SSQ_r$ met $SSQ_r := \|V(B\underline{b}^0 - \delta)\|^2$.

En evenals 7.1: $SSQ_r \sim \sigma^2 \chi^2_{q}$ met $\delta_r = \|V(B\beta - \delta)\|^2 = 0$, nl. $\mathcal{E}B\underline{b}^0 = BI^0\beta = B\beta$ daar $B\beta$ schatbaar is.

Hieruit volgt: $\mathcal{E}(SSE_r) = (n - r + q)\sigma^2$ daar $\mathcal{E}\chi^2_\nu = \nu$ zodat

11.3. $$\underline{s}^2 = \frac{SSE_r}{n - r + q}$$

een zuivere schatter is voor $\sigma^2$ ($r$ = rang van $X$).

De schatbaarheid van een functie $k'\beta$ in het niet beperkte model blijft behouden in het beperkte model. De voorwaarde voor de schatbaarheid, nl. - er is een $t$ zodat $\mathcal{E}(t'\underline{y}) = k'\beta$ - blijft gehandhaafd. Echter, ofschoon de functie schatbaar blijft, is het een functie van de parameters en is onderworpen aan de beperkingen $B\beta = \delta$. De vorm van $k'\beta$ kan veranderen.

In ons voorbeeld is $k'\beta = \mu + \frac{1}{2}(\alpha_1 + \alpha_2)$ een schatbare functie in het niet beperkte model.

Beperken we ons model door $\alpha_1 - \alpha_2 = 0$, dan wordt $k'\beta$ gelijk aan $\mu + \alpha_1$ of $\mu + \alpha_2$.

In het algemeen verandert een schatbare functie $k'\beta$ in het niet beperkte model in $k'\beta + \lambda'(B\beta - \delta)$ in het beperkte model, waarbij $\lambda'$ geschikt gekozen wordt en wel zō dat $\lambda'\delta = 0$ omdat het anders geen functie van de parameters alleen blijft. Als $\delta = 0$ is iedere $\lambda$ toelaatbaar. Als $q = 1$ is $\delta$ een scalar en $\lambda = 0$; dan blijft $k'\beta$ in het beperkte model dezelfde gedaante behouden als in het niet beperkte.

Toetsbare hypothesen $K\beta = \gamma$ in het niet beperkte model blijven toetsbaar in het beperkte model, <u>mits</u> de vergelijkingen $K\beta = \gamma$ en $B\beta = \delta$ samen een consistent stelsel vormen en weer $\langle K' \rangle \cap \langle B' \rangle = \{0\}$.

ii) De elementen van $B\beta$ zijn niet schatbaar.

De situatie is nu geheel analoog aan die van een niet toetsbare hypothese $K\beta = \gamma$ en er is dan ook een oplossing van het oorspronkelijke stelsel normaalvergelijkingen in het <u>niet beperkte</u> model die tevens voldoet aan de restricties.

11.4. Deze oplossing is: $\underline{b}_r^0 = \underline{b}^0 + (I^0 - I)\underline{z}_1$ waarin weer $\underline{z}_1$ voldoet aan

$$B(I^0 - I)z_1 = \delta - BGX'\underline{y} .$$

Er geldt dus dat $X'X\underline{b}_r^0 = X'\underline{y}$ en $B\underline{b}_r^0 = \delta$ zodat $SSE_r = SSE$ oftewel $SSQ_r = 0$.

Een schatbare functie in het niet beperkte model blijft schatbaar in het beperkte model. Door de beperkingen wordt de gedaante wel veranderd. Omdat evenwel de beperkingen geen schatbare functies zijn is het mogelijk dat de aangepaste vorm van een schatbare functie in het beperkte model wel en in het niet beperkte model niet schatbaar is.

In ons voorbeeld is $\mu + \frac{1}{2}(\alpha_1 + \alpha_2)$ een schatbare functie in het niet beperkte model. In het door $\alpha_1 = 0$ beperkte model wordt dit $\mu + \frac{1}{2}\alpha_2$ wat wel in het beperkte model doch niet in het niet beperkte model een schatbare functie is.

b. <u>Gebruikelijke beperkingen op de oplossingen</u>

De bron van alle moeilijkheden in het niet-volle rang model is het feit dat $r(X) < p$ en dus $X'X$ singulier is. Tot nu toe is dat ondervangen door het gebruik van de gegeneraliseerde inverse. Een andere aanpak om een ēēnduidige oplossing te verkrijgen is het opleggen van "gebruikelijke beperkingen" (usual constraints) aan de <u>oplossingen</u> van de normaalvergelijkingen: $X'X\underline{b} = X'\underline{y}$.

Een voorbeeld: Stel

$$X = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix};$$

de normaalvergelijkingen worden

$$\begin{aligned} 6\mu^0 + 2a_1^0 + 2a_2^0 + 2a_3^0 &= y_{..} \\ 2\mu^0 + 2a_1^0 \phantom{+ 2a_2^0 + 2a_3^0} &= y_{1.} \\ 2\mu^0 \phantom{+ 2a_1^0} + 2a_2^0 \phantom{+ 2a_3^0} &= y_{2.} \\ 2\mu^0 \phantom{+ 2a_1^0 + 2a_2^0} + 2a_3^0 &= y_{3.} \end{aligned}$$

Leggen we aan de oplossing de beperking op $a_1^0 + a_2^0 + a_3^0 = 0$, dan zien we gemakkelijk in dat er onder deze beperking één oplossing is nl.

$$(\mu^0, a_1^0, a_2^0, a_3^0) = (\bar{y}_{..}, \bar{y}_{1.} - \bar{y}_{..}, \bar{y}_{2.} - \bar{y}_{..}, \bar{y}_{3.} - \bar{y}_{..}) .$$

Elk stel gebruikelijke beperkingen op de oplossingen dat één oplossing geeft kan worden gekozen. In orthogonale schema's zijn deze van de vorm $\Sigma a_i^0 = 0$ vaak voor de hand liggend, ze komen dan ook vaak overeen met de beperkingen in het model zodat dan eigenlijk een model van volle rang is ontstaan.
(In niet orthogonale schema's is dit type beperkingen aan de oplossingen in het algemeen niet het eenvoudigste). Dit geldt trouwens steeds: als beperkingen op de oplossingen, waardoor een éénduidige oplossing ontstaat, overeenkomen met beperkingen in het model aan de parameters, ontstaat een model van volle rang en zijn de oplossingen van de normaalvergelijkingen van het beperkte model ook schattingen voor de parameters in het beperkte model.
We hebben gezien dat met een willekeurige $\underline{b}^0$, oplossing van de normaalvergelijkingen, veel van de in een lineair model belangrijke grootheden gevonden kunnen worden, zoals: SSE, de BLUE van $k'\beta$ e.d. Deze kunnen bepaald worden hoe ook de $\underline{b}^0$ gevonden is.
Om de schatbaarheid van een functie $k'\beta$, de toetsbaarheid van een hypothese $K\beta = \gamma$ te onderzoeken hebben we evenwel de bij de $\underline{b}^0$ behorende gegeneraliseerde inverse nodig. Ofschoon het opleggen van beperkingen aan de oplossingen van de normaalvergelijkingen vaak de eenvoudigste weg geeft om een op-

lossing te vinden, moeten we voor bepaalde doeleinden (schatbaarheid, e.d.) dus toch een gegeneraliseerde inverse van X'X vinden die bij deze oplossing behoort. Immers de gebruikelijke beperkingen aan de oplossingen worden uitsluitend gebruikt om een oplossing te vinden. Deze eis dwingt tot nadenken over de keuze van de gebruikelijke beperkingen om zo mogelijk tot een oplossing te komen waarbij een gegeneraliseerde inverse van X'X met niet al te veel problemen gevonden kan worden.

Voorwaarden aan de gebruikelijke beperkingen

Benadrukt wordt nogmaals dat de beperkingen aan de oplossingen alleen toegepast worden bij het oplossen van de normaalvergelijkingen $X'X\underline{b} = X'\underline{y}$ van het niet beperkte model, om een éénduidige oplossing te vinden en dat ze niet noodzakelijk iets uit te staan hebben met de modelparameters.
Laat de beperkingen gegeven worden door $Cb^0 = \gamma$, dan wordt het stelsel normaalvergelijkingen aangevuld tot

11.5. $$X'Xb^0 = X'y$$
$$Cb^0 = \gamma,$$

hetgeen we kunnen herschrijven tot

$$\begin{pmatrix} X'X \\ C \end{pmatrix} b^0 = \begin{pmatrix} X'y \\ \gamma \end{pmatrix}$$

$$X'X \in \mathfrak{M}^r_{p,p} .$$

Waaraan moet C voldoen opdat 11.4 een éénduidige oplossing heeft?
Van de eerste p vergelijkingen (die consistent zijn) schrappen we er p - r, zodat er r onafhankelijke overblijven ($r = r(X'X)$). Wil het nu ontstane stelsel precies 1 oplossing hebben, dan is dus dim oplossingsvariëtiet 0, dan moeten er p onafhankelijke vergelijkingen zijn in de p onbekende componenten van $b^0$.
Nu kunnen we altijd de r onafhankelijke rijen van X'X aanvullen met p - r onafhankelijke rijen $c'_{i*}$ zodat er p onafhankelijke vergelijkingen ontstaan; $\gamma$ is dan nog volkomen willekeurig, het stelsel is toch consistent.

De matrix C voldoet precies aan de voorwaarde die in de definitie van een niet toetsbare hypothese aan K gesteld is, nl. $\langle C' \rangle \cap \langle X'X \rangle = \{0\}$ en $r(C) = p - r$. Beschouwen we vervolgens de matrix

$$S := \begin{pmatrix} X'X & C' \\ C & 0 \end{pmatrix}.$$

Deze is regulier (zie 1.55). In D11 zagen we dat, als

$$S^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}, \quad B_{11} = (X'X)^- .$$

De oplossing $b^0$ van 11.5 kan ook nog anders verkregen worden, nl. analoog
11.6. aan 11.4 door $\underline{b}^0 = GX'\underline{y} + (I^0 - I)\underline{z}_1$ waarbij $z_1$ de oplossing is van het stelsel

$$C(I^0 - I)\underline{z}_1 = \gamma - CGX'\underline{y} .$$

Opgemerkt is reeds dat voor bepaalde doeleinden een $(X'X)^-$ die bij de oplossing $\underline{b}^0$ van 11.5 behoort, nodig is. Dit is in het algemeen niet de matrix G in 11.6. G is meestal een eenvoudig te vinden $(X'X)^-$.
Met 1.43 kunnen we schrijven $\underline{z}_1 = ZX'\underline{y}$. Hiermee wordt 11.6

$$\underline{b}^0 = [GX' + (I^0 - I)ZX']\underline{y} .$$

De matrix $G + (I^0 - I)Z$ is een $(X'X)^-$ die behoort bij de oplossing $\underline{b}^0$ van het stelsel 11.5 (zie D14).

In het geval dat $\gamma = 0$, geldt dat een bij de oplossing $\underline{b}^0$ van de vergelijkingen 11.5 behorende $(X'X)^-$ de hiervoor genoemde $B_{11}$ is. Immers, omdat

$$\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

de inverse van

$$\begin{pmatrix} X'X & C' \\ C & 0 \end{pmatrix}$$

is geldt dat $CB_{11} = 0$ en dus met $\underline{b}^0 = B_{11}X'\underline{y}$ dat $C\underline{b}^0 = 0$, zodat inderdaad $\underline{b}^0$ als oplossing van 11.5 en $B_{11}$ bij elkaar horen.

Beperkingen van de vorm $b_i^0 = 0$

In niet orthogonale schema's zijn de eenvoudigste beperkingen aan de oplossingen het "nul" stellen van p - r van de componenten van $b^0$. Vanzelfsprekend kan niet elk willekeurig (p - r)-tal worden genomen daar de matrix C die deze beperkingen weergeeft, moet voldoen aan de bovengenoemde voorwaarden.

Ons voorbeeld 1

Het stelsel normaalvergelijkingen X'Xb = X'y is nu:

$$\begin{aligned}
6\mu^0 + 3a_1^0 + 2a_2^0 + a_3^0 &= 96 \\
3\mu^0 + 3a_1^0 \qquad\qquad\quad &= 45 \\
2\mu^0 \qquad\quad + 2a_2^0 \qquad &= 24 \\
\mu^0 \qquad\qquad\qquad\quad + a_3^0 &= 27\ .
\end{aligned}$$

De aanvullende bepaling is $a_1^0 + a_2^0 + a_3^0 = 0$.

1e methode: (direct oplossen van het stelsel 11.5)

$$\begin{pmatrix} 6 & 3 & 2 & 1 \\ 3 & 3 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \mu^0 \\ a_1^0 \\ a_2^0 \\ a_3^0 \end{pmatrix} = \begin{pmatrix} 96 \\ 45 \\ 24 \\ 27 \\ 0 \end{pmatrix}$$

waaruit het stelsel

$$\underbrace{\begin{pmatrix} 3 & 3 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}}_{A} \begin{pmatrix} \mu^0 \\ a_1^0 \\ a_2^0 \\ a_3^0 \end{pmatrix} = \begin{pmatrix} 45 \\ 24 \\ 27 \\ 0 \end{pmatrix}$$

te voorschijn komt.

De oplossing is

$$\begin{pmatrix} \mu^0 \\ a_1^0 \\ a_2^0 \\ a_3^0 \end{pmatrix} = \frac{1}{18} \underbrace{\begin{pmatrix} 2 & 3 & 6 & -6 \\ 4 & -3 & -6 & 6 \\ -2 & 6 & -6 & 6 \\ -2 & -3 & 12 & 6 \end{pmatrix}}_{B} \begin{pmatrix} 45 \\ 24 \\ 27 \\ 0 \end{pmatrix} = \begin{pmatrix} 18 \\ -3 \\ -6 \\ 9 \end{pmatrix}$$

waarin $B = A^{-1}$.

<u>2e methode</u>: (met 11.6 om <u>dezelfde</u> éénduidig bepaalde oplossing $b^0$ te vinden)
De matrix G uit 11.6 kan eenvoudig genomen worden; we hadden reeds

$$G = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{en} \quad GX'y = \begin{pmatrix} 0 \\ 15 \\ 12 \\ 27 \end{pmatrix} .$$

Met $C = (0 \quad 1 \quad 1 \quad 1)$ worden deze

$$(0 \quad 1 \quad 1 \quad 1) \begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} z_{11} \\ z_{12} \\ z_{13} \\ z_{14} \end{pmatrix} = 0 - (0 \quad 1 \quad 1 \quad 1) \begin{pmatrix} 0 \\ 15 \\ 12 \\ 27 \end{pmatrix} .$$

(Voor $I^0$ zie § 8g).

<u>Opmerking</u>: Alleen $z_{11}$ is de variabele die een rol speelt ($p - r$ variabelen; $p = 4$, $r = 3$) $3z_{11} = -54$; $z_{11} = -18$.
Dus elke $\underline{z}_1' = (z_{11}, z_{12}, z_{13}, z_{14})$ is goed mits $z_{11} = -18$.
De oplossing $b^0$ wordt

$$b^0 = \begin{pmatrix} \mu^0 \\ a_1^0 \\ a_2^0 \\ a_3^0 \end{pmatrix} = \begin{pmatrix} 0 \\ 15 \\ 12 \\ 27 \end{pmatrix} + \begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & & & \\ 1 & & O & \\ 1 & & & \end{pmatrix} \begin{pmatrix} -18 \\ z_{12} \\ z_{13} \\ z_{14} \end{pmatrix} = \begin{pmatrix} 18 \\ -3 \\ -6 \\ 9 \end{pmatrix} .$$

<u>Ons voorbeeld 2</u>

Laten we tot slot voorbeeld 2 van pagina 8.3 eens geheel uitwerken.
Stel de waarnemingen zijn:

| | | 1 | 2 | kunstmest |
|---|---|---|---|---|
| | 1 | 3;4 | 11 | |
| ras | 2 | 1 | 3 | |
| | 3 | 8 | | |

model: $y_{ijk} = \mu + \alpha_i + \beta_j + \underline{e}_{ijk}$ oftewel $\underline{y} = X\beta + \underline{e}$ met $y = (3,4,11,1,3,8)'$
en $\beta = (\mu,\alpha_1,\alpha_2,\alpha_3,\beta_1,\beta_2)'$. (gekruist, 2 factorenmodel zonder interactie)

$$X = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \end{pmatrix} \; ; \; X'X = \begin{pmatrix} 6 & 3 & 2 & 1 & 4 & 2 \\ & 3 & 0 & 0 & 2 & 1 \\ & & 2 & 0 & 1 & 1 \\ & & & 1 & 1 & 0 \\ & * & & & 4 & 0 \\ & & & & & 2 \end{pmatrix}$$

$$X'y = (30,18,4,8,16,14)'$$

$$G = \frac{1}{7}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 2 & -2 & -3 \\ 0 & 0 & 2 & 10 & -3 & -1 \\ 0 & 0 & -2 & -3 & 3 & +1 \\ 0 & 0 & -3 & -1 & 1 & 5 \end{pmatrix} \; ; \; b^0 = GX'y = \frac{1}{7}(0,0,-34,26,30,66)'$$

$$\hat{y} = Xb^0 = \frac{1}{7}(30,30,66,-4,32,56)' \; .$$

(Controleer $\Sigma\hat{y}_i = \Sigma y_i = 30$); $r(X) = 4$.

$$I^0 = GX'X = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & & & & \\ 0 & -1 & & & & \\ 1 & 1 & & I_4 & & \\ 1 & 1 & & & & \end{pmatrix} ;$$

Stel we willen toetsen $H_0$: er is geen verschil in ras oftewel $\alpha_1 = \alpha_2 = \alpha_3$: $K\beta = 0$ met

$$K = \begin{pmatrix} 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \end{pmatrix} ;$$

Deze $H_0$ is toetsbaar daar $KI^0 = K$ (ga na).

$V^2 = (KGK')^{-1} = \frac{1}{4}\binom{6\;2}{2\;3}$; $b_0'K' = \frac{1}{7}(34,-60)$; SSQ = 49; $r(K) = q = 2$. SSE = 9,2.

$$F(Q) = F_2^2 = \frac{49/2}{9{,}2/2} = 5{,}3 \quad \text{niet significant.}$$

Opmerking: in zo'n model worden in het algemeen de normaalvergelijkingen:

$$\left(\begin{array}{c|ccc|cc} n_{\cdot\cdot} & n_{1\cdot} & n_{2\cdot} & n_{3\cdot} & n_{\cdot 1} & n_{\cdot 2} \\ \cdot\cdot & \cdots & \cdots & \cdots & \cdots & \cdots \\ & n_{1\cdot} & 0 & & n_{11} & n_{12} \\ & & n_{2\cdot} & & n_{21} & n_{22} \\ & & & n_{3\cdot} & n_{31} & n_{32} \\ \hline & & & & n_{\cdot 1} & 0 \\ & & & & & n_{\cdot 2} \end{array}\right) \begin{pmatrix} \mu \\ \hline \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \hline \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} y_{\cdot\cdot\cdot} \\ \hline y_{1\cdot\cdot} \\ y_{2\cdot\cdot} \\ y_{3\cdot\cdot} \\ \hline y_{\cdot 1\cdot} \\ y_{\cdot 2\cdot} \end{pmatrix}$$

$n_{ij}$ = aantal waarnemingen in cel ij, dus voor ras i en kunstmest j.

## § 12. Het algemene Gauss-Markoff model

12.1. Definitie: Het tripel $(y, X\beta, \sigma^2 V)$ is het GGM model (= general Gauss-Markoff model), d.w.z.:
$\underline{y} = X\beta + \underline{e}$; $\mathcal{E}\underline{e} = 0$; var $\underline{e} = \sigma^2 V$ ($\sigma^2$, $\beta$ onbekend; V bekend) $X_{n,p}$ en V mogen singulier zijn.

Hierin is bevat het model met lineaire beperkingen (constraints) $B\beta = \delta$ op de parameters.

Notatie: $(y, X\beta \mid B\beta = \delta, \sigma^2 V)$ nl.:

Definieer: $y_r' = (y' \mid \delta')$; $X_r' = (X' \mid B')$ en $V_r = \begin{pmatrix} V0 \\ 00 \end{pmatrix}$.

Dit geeft:

$$\binom{y}{\delta} = \binom{X}{B}\beta + \binom{e}{0} \quad \text{oftewel } (y_r, X_r\beta, \sigma^2 V_r).$$

12.2. Definitie: Het tripel $(y, X\beta, \sigma^2 I)$ met $r = r(X) = p$, dus van volle rang, heet het SGM model (= standaard Gauss-Markoff model).

12.3. Bewering: Een willekeurig GGM model is altijd terug te voeren tot een SGM model. In zekere zin ontstaat er dus geen nieuw probleem. We laten dit stapsgewijs zien.

a) Onderstel $(y, X\beta, \sigma^2 V)$ met X, V van volle rang.
$V = D^2$, D regulier, symmetrisch. Stel $y_* = D^{-1}y$; $X_* = D^{-1}X$ dan wordt $\underline{y} = X\beta + \underline{e} : D^{-1}\underline{y} = D^{-1}X\beta + D^{-1}\underline{e}$ oftewel $(y_*, X_*\beta, \sigma^2 I)$.

b) Onderstel $(y, X\beta, \sigma^2 I)$ met $r = r(X) < p$.
Volgens (1.81) is $X = P\Lambda^{\frac{1}{2}}Q'$, $P^2 = Q^2 = I_r$ dus $r(P) = r$, volle rang.
$y = X\beta$ wordt $y = P\Lambda^{\frac{1}{2}}Q'\beta =: P\beta_*$, zodat ontstaat een SGM model $(y, P\beta_*, \sigma^2 I)$.

Opmerking: de normaalvergelijkingen $X'X\hat{\beta} = X'y$ worden $\hat{\beta}_* = P'y$.
en

$$s^2 = \frac{y'y - \hat{y}'\hat{y}}{n-r} \text{ met } \hat{y} = P\hat{\beta}_* = PP'y \text{ dus } \hat{y}'\hat{y} = \hat{\beta}_*'\hat{\beta}_*.$$

c) Onderstel $(y, X\beta, \sigma^2 V)$ met $r = r(X) < p$, $V > 0$.

Dit doen we in twee stappen. Terugbrengen tot a) via $X_* = D^{-1}X$ geeft $(y_*, X_*\beta, \sigma^2 I)$.
Dan volgens b) verder met $X_*$ i.p.v. X. Dat geeft $(y_*, P_*\beta_*, \sigma^2 I)$, een SGM model.

d) Onderstel $(y, X\beta | B\beta = \delta, \sigma^2 I)$ met $r\binom{X}{B} = p$, dus volle rang.
Vanwege de constraints $B\beta = \delta$ kunnen we enkele parameters elimineren en ontstaat een <u>gereduceerd</u> model $(y, X_1\beta_1, \sigma^2 I)$, een SGM model.
Is $r\binom{X}{B}$ of $r(X_1)$ niet vol, dan verder als onder b).

e) Onderstel nu <u>algemeen</u> $(y, X\beta, \sigma^2 V)$ met $r = r(X) \le p$ en $r(V) = s \le n$.
Beschouw de eigenwaarden en eigenvectoren van V:

$$\begin{cases} \lambda_1^2,\ldots,\lambda_s^2 \text{ alle} > 0 \text{ en } & 0,\ldots,0 \\ f_1,\ldots,f_s & g_1,\ldots,g_{n-s} \end{cases}$$

$$F_{n,s} := (f_1/\lambda_1 \;\ldots\; f_s/\lambda_s) \quad \text{en} \quad G_{n,n-s} := (g_1 \;\ldots\; g_{n-s})$$

Transformeer

$$\begin{matrix} y_1 = F'y \\ y_2 = G'y \end{matrix} \quad \text{geeft} \begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \end{pmatrix} = \begin{pmatrix} F' \\ G' \end{pmatrix} \underline{y} = \begin{pmatrix} F' \\ G' \end{pmatrix} X\beta + \underline{e}_* = \begin{pmatrix} F'X \\ G'X \end{pmatrix}\beta + \underline{e}_* =: \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}\beta + \underline{e}_*$$

Dit is juist het model $(y_1, X_1\beta | X_2\beta = y_2, \sigma^2 I)$ nl.:
var $\underline{y}_2 = \sigma^2 G'VG = 0$ daar $VG = 0$, dus $\underline{y}_2$ is constant $\delta$.
En var $\underline{y}_1 = \sigma^2 F'VF = \sigma^2 I$ indien eigenvectoren georthonormaliseerd zijn!
Dus teruggebracht tot geval d).

## 12.4. <u>Consistentie van het GGM model</u>

1) var $\underline{y} = \sigma^2 V$. Is V singulier, dan is er een vector $\alpha$ zodat $\alpha'\underline{y}$ = constant.
Nu volgt uit $\alpha'X = 0$ èn $\alpha'V = 0$, dat ook $\alpha'\underline{y} = 0$ nl.:
$\underline{y} = X\beta + \underline{e}$; $\mathcal{E}\alpha'\underline{y} = \alpha'X\beta = 0$ en var $\alpha'\underline{y} = \alpha'V\alpha = 0$ d.w.z. $\alpha'\underline{y} = 0$.
Dus $\alpha \in \langle X \mid V\rangle^\perp \Rightarrow \alpha \in \langle y\rangle^\perp$. Gevolg $\langle y\rangle^\perp \supset \langle X \mid V\rangle^\perp$
oftewel $\underline{y} \subset \langle X \mid V\rangle$ (met kans 1).

2) $\alpha'V = 0$ alleen geeft $\alpha'(\underline{y} - X\beta) = 0$ nl. $\mathcal{E}(\alpha'(\underline{y} - X\beta)) = \alpha'0 = 0$
en var $\alpha'(\underline{y} - X\beta) = \alpha'V\alpha = 0$ dus $\underline{y} - X\beta \in \langle V\rangle$ (met kans 1).

Opmerking: Is aan 1) en 2) voldaan, dan heet het model consistent.
Is V regulier, dan is vanzelf aan beide eisen voldaan. Singulariteit van V legt dus beperkingen op aan y en $\beta$.

## 12.5. Basistheorema

Beschouw het consistente GGM model $(y, X\beta, \sigma^2 V)$
Zij

$$\begin{pmatrix} V & X \\ X' & 0 \end{pmatrix}^- = \begin{pmatrix} C_1 & C_2 \\ C_3 & -C_4 \end{pmatrix}; \quad H_0 : K\beta = \gamma$$

een toetsbare hypothese dan geldt:

a) $K\underline{b}^\circ = KC_2'\underline{y}$ bevat de blues voor $K\beta$.

b) $\mathrm{var}(K\underline{b}^\circ) = \sigma^2 KC_4K'$.

c) Een zuivere schatter voor $\sigma^2$ is $\underline{s}^2 = \underline{y}'C_1\underline{y}/\nu$ met $\nu = r(V \mid X) - r(X)$.

Bewijs:

a) $K = TX$; $\mathcal{E}(K\underline{b}^\circ) = \mathcal{E}(TXC_2'\underline{y}) = TXC_2'X\beta = TX\beta = K\beta$ (zie D 21 a).
Stel $q'\underline{b}^\circ = \ell'\underline{y}$ dus $\mathcal{E}(\ell'\underline{y}) = \ell'X\beta = q'\beta$ dus $q = X'\ell$
èn $\mathrm{var}(\ell'\underline{y}) = \ell'V\ell$ minimaal.
Dus een oplossing gezocht van $X'\ell = q$ en $\|\ell\|_V^2$ minimaal,
d.w.z. (zie D21 a) $\ell = (X')^-_{m(V)}q = C_2\, q$
Dus $\ell'\underline{y} = q'C_2'\underline{y}$ is BLUE voor $q'\beta$.

b) $\mathrm{var}(K\underline{b}^\circ) = \mathrm{var}(TXC_2'\underline{y}) = \sigma^2(TXC_2'VC_2X'T') = \sigma^2(TXC_4X'C_2X'T') =$

$$= \sigma^2(TXC_4X'T') = \sigma^2(KC_4K') \quad \text{(met D 21 a,c)}.$$

Volgens D 21 c is dit een symmetrische matrix, daar $XC_4X'$ symmetrisch is.

c) $\mathcal{E}(\nu \underline{s}^2) = \mathcal{E}(\underline{y}'C_1\underline{y}) \overset{*}{=} \sigma^2\ \mathrm{sp}(C_1V) = [r(V \mid X) - r(X)]\ \sigma^2 = \nu\sigma^2$. (D21 e).

In $*$ is (2.36) gebruikt met $\mu'A\mu = 0$ nl. $(X\beta)'C_1X\beta = \beta'X'C_1X\beta = 0$ (D21 b).

Opmerking: Is $V = I$, zoals in § 5 t/m § 11 aangenomen was, dan komen de bekende relaties zoals 10.3 terug. (zie opgaaf 12.1).

12.6. Stelling: $\underline{y} \sim N_n(X\beta, \sigma^2 V)$, $K\underline{b}^\circ$ BLUE voor $K\beta$ $(K = TX)$.

Er geldt:

a) $K\underline{b}^\circ \sim N_q\ (K\beta, \sigma^2 KC_4K')$ onafhankelijk van

$$\nu \underline{s}^2/\sigma^2 \sim \chi^2_\nu \quad \text{met } \nu = r(V \vdots X) - r(X)$$

b) Zij $H_0 : K\beta = \gamma$. Deze is consistent d.e.s.d. als

$$DD^-(Kb^\circ - \gamma) = Kb^\circ - \gamma \quad \text{met } D := KC_4K'.$$

Indien consistent, dan geldt onder $H_0$:

$$(K\underline{b}^\circ - \gamma)(KC_4K')^-(K\underline{b}^\circ - \gamma)/d\underline{s}^2 \simeq F^d_\nu \qquad \text{met } d = r(KC_4K').$$

Bewijs:

a) De verwachting en variantie van $K\underline{b}^\circ$ zijn reeds bepaald onder 12.5.

De normaliteit van $K\underline{b}^\circ = KC_2'\underline{y}$ volgt uit de normaliteit van $\underline{y}$. □

Uit (4.11) volgt:

$$\nu \underline{s}^2/\sigma^2 = \underline{y}'C_1\underline{y}/\sigma^2 \simeq \underline{\chi}^2_\nu \leftrightarrow VC_1VC_1V = VC_1V \text{ en } \nu = \mathrm{sp}(C_1V) = r(V \vdots X) - r(X)$$

Als $C_1$ niet symmetrisch is, is voldoende dat $VC_1V$ symmetrisch is.

Dit volgt uit D 21 d en g. $\mu'A\mu$ wordt $\beta'X'C_1X\beta = 0$, dus centrale $\chi^2_\nu$. □

Nu de onafhankelijkheid:

$$K\underline{b}^\circ = KC_2'\underline{y} = TXC_2'\underline{y} \text{ en } \underline{y}'C_1\underline{y} \text{ zijn o.o. volgens } (4.14^*)$$

d.e.s.d. als $TXC_2'VC_1V = 0$. Dit is zo nl. $= TVC_2X'C_1V = 0$ (D 21 a,b) □

Tevens is $TXC_2'VC_1X\beta = 0$.

b) $(K\underline{b}^\circ - \gamma)D^-(K\underline{b}^\circ - \gamma)/\sigma^2 \simeq \chi^2 d$ d.e.s.d. volgens 4.11* als
$DD^-DD^-D = DD^-D$ (dit is zo) èn $DD^-D$ symmetrisch (= D, is symmetrisch, zie 12.5 b) èn $d = sp(D^-D) = r(D^-D) = r(D) = d$.
Onder $H_0$ is $\mathcal{E}K\underline{b}^\circ = K\beta = \gamma$ dus centrale $\chi^2$.
Daar teller en noemer o.o., ontstaat een centrale F-stochastiek. □
Consistentie van $H_0 : K\beta = \gamma$:
Onder $H_0$ is $\mathcal{E}[\ell'(K\underline{b}^\circ - \gamma)] = 0$. Dus consistent als $\forall \ell$ met var $\ell'(K\underline{b}^\circ - \gamma) = 0$ volgt $\ell'(K\underline{b}^\circ - \gamma) = 0$. Oftewel $\ell'D\ell = 0 \Rightarrow \ell'(K\underline{b}^\circ - \gamma) = 0$
$\ell'D\ell \Rightarrow \ell'D = 0$ dus $\ell \perp \langle D\rangle \Rightarrow \ell \perp K\underline{b}^\circ - \gamma$ oftewel $\langle D\rangle^\perp \subset \langle K\underline{b}^\circ - \gamma\rangle^\perp$.
Dus $K\underline{b}^\circ - \gamma \in \langle D\rangle$ en dit is equivalent met $DD^-(K\underline{b}^\circ - \gamma) = K\underline{b}^\circ - \gamma$ (zie som D 15).

<u>Opmerking</u>: $\underline{y}'C_1\underline{y} = (\underline{y} - X\beta)'C_1(\underline{y} - X\beta)$ nl.:

$$= \underline{y}'C_1\underline{y} - \beta'X'C_1\underline{y} - \underline{y}'C_1X\beta + \beta'X'C_1X\beta$$

Nu is laatste term 0 volgens D 21 b.

$$\mathcal{E}(\beta'X'C_1\underline{y}) = \beta'X'C_1X\beta = 0 \text{ èn var } (\beta'X'C_1\underline{y}) = \beta'X'C_1VC_1'X\beta = 0 \quad \text{(D 21 b)}.$$

Dus $\beta'X'C_1\underline{y} = 0$. Analoog $\underline{y}'C_1X\beta = 0$. □

## Appendix A: Normaal verdeelde, doch niet simultaan normaal verdeelde variabelen.

Stel $\underline{x}$ en $\underline{y}$ hebben als simultane kansverdeling (voorbeeld J. van IJzeren)

$$f(x,y) = \frac{1}{2\pi\sqrt{2}}(e^{-x^2-xy\sqrt{2}-y^2} + e^{-x^2+xy\sqrt{2}-y^2}), \qquad -\infty < x,y < \infty .$$

1) Er geldt: zijn $\underline{x}$ en $\underline{y}$ *simultaan* normaal verdeeld, dan zijn $\underline{x}$ en $\underline{y}$ ook elk normaal verdeeld. Echter het omgekeerde geldt niet. Stel nl. $\underline{x}$ en $\underline{y}$ als boven verdeeld. Dan is

$$f_x(x) = \frac{1}{2\pi\sqrt{2}} e^{-\frac{1}{2}x^2}\left[\int_{-\infty}^{\infty} e^{-(y+\frac{1}{2}x\sqrt{2})^2}dy + \int_{-\infty}^{\infty} e^{-(y-\frac{1}{2}x\sqrt{2})^2}dy\right] = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} .$$

Analoog $f_y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}$. Dus $\underline{x}$, $\underline{y}$ elk normaal verdeeld, doch niet simultaan normaal.

2) Er geldt: zijn $\underline{x}$ en $\underline{y}$ *simultaan* normaal verdeeld met $\rho(\underline{x},\underline{y}) = 0$, dan zijn $\underline{x}$ en $\underline{y}$ stochastisch onafhankelijk. Zijn $\underline{x}$ en $\underline{y}$ elk normaal verdeeld, doch niet simultaan normaal verdeeld, dan is dit niet juist; nl. neem $\underline{x}$ en $\underline{y}$ weer als boven verdeeld:

$$\rho(\underline{x},\underline{y}) = \iint xy\ f(x,y)dxdy .$$

Nu is

$$\int y\ f(x,y)dy = \frac{1}{2\pi\sqrt{2}} e^{-\frac{1}{2}x^2}\left[\int_{-\infty}^{\infty} y\ e^{-(y+\frac{1}{2}x\sqrt{2})^2}dy + \int_{-\infty}^{\infty} y^{-(y-\frac{1}{2}x\sqrt{2})^2}dy\right] =$$

$$= \frac{1}{2\pi\sqrt{2}} e^{-\frac{1}{2}x^2}\left[\int_{-\infty}^{\infty}(u - \tfrac{1}{2}x\sqrt{2})e^{-u^2}du + \int_{-\infty}^{\infty}(v + \tfrac{1}{2}x\sqrt{2})e^{-v^2}dv\right] = 0 ,$$

dus $\rho(\underline{x},\underline{y}) = 0$.

Toch zijn $\underline{x}$ en $\underline{y}$ stochastisch afhankelijk daar $f(x,y) \neq f_x(x)f_y(y)$.

3) Er geldt: de som van *simultaan* normaal verdeelde variabelen is normaal verdeeld. Dit geldt niet als we het woord "simultaan" weglaten. Stel weer $\underline{x}$ en $\underline{y}$ als boven verdeeld. Voer de volgende transformatie uit: $\underline{x} = (\underline{u} + \underline{v})/\sqrt{2}$ en $\underline{y} = (\underline{u} - \underline{v})/\sqrt{2}$
We zoeken de kansdichtheid van $\underline{x} + \underline{y} = \underline{u}\sqrt{2}$. Dan is (det. v. Jacobi = 1)

$$g(u,v) = \frac{1}{2\pi\sqrt{2}} \{\exp[-u^2(1 + \tfrac{1}{2}\sqrt{2}) - v^2(1 - \tfrac{1}{2}\sqrt{2})] + \exp[-u^2(1 - \tfrac{1}{2}\sqrt{2}) - v^2(1 + \tfrac{1}{2}\sqrt{2})]\}$$

en dus

$$g_u(u) = \int g(u,v)dv = \frac{1}{2\sqrt{\pi}}\left[\frac{1}{\sqrt{2 - \sqrt{2}}}\, e^{-u^2(1+\frac{1}{2}\sqrt{2})} + \frac{1}{\sqrt{2 + \sqrt{2}}}\, e^{-u^2(1-\frac{1}{2}\sqrt{2})}\right]$$

d.w.z. $\underline{x} + \underline{y}$ niet normaal verdeeld!

## Appendix B. Differentiatie van scalaire functies van de elementen van een matrix

a. Is $y = f(x_1,\dots,x_n)$ een minstens tweemaal differentieerbare scalaire functie van $x_1,\dots,x_n$, dan is bekend uit de theorie van functies van meerdere variabelen:

$$dy = \sum_i \frac{\partial y}{\partial x_i}\, dx_i \quad \text{en} \quad d^2y = \sum_{i,j} \frac{\partial^2 y}{\partial x_i\, \partial y_j}\, dx_i dy_j \;.$$

Voldoende voor het optreden van een maximum in een punt $P_0$ is dat $(dy)_{P_0} = 0$ en $(d^2y)_{P_0} < 0$.

Zij nu $y = y(X)$ een minstens tweemaal differentieerbare scalaire functie van de elementen van een matrix $X_{n\times p}$.

Definitie 1: $\frac{dy}{dX} := \{\frac{\partial y}{\partial x_{ji}}\}$ (dus een matrix van de orde $p \times n$). $\frac{\partial y}{\partial X'} = (\frac{\partial y}{\partial X})'$,

$i = 1,\dots,n;\ j = 1,\dots,p.$

Definitie 2: $dX := \{dx_{ij}\}$.

b. Eigenschappen:

B1. $dy = sp(\frac{\partial y}{\partial X}\, dX)$. Eveneens als $y = y(X(t))$ : $\frac{dy}{dt} = sp(\frac{\partial y}{\partial X}\frac{dX}{dt})$.

Bewijs: $y = f(x_{11},\dots,x_{np})$ dus

$$dy = \sum_{i,k} \frac{\partial y}{\partial x_{ik}}\, dx_{ik} = \sum_k \sum_i (\frac{\partial y}{\partial X})_{ki} (dX)_{ik} = \sum_k (\frac{\partial y}{\partial X}\, dX)_{kk} = sp(\frac{\partial y}{\partial X}\, dX) \;.$$

Opmerking: $dX' = (dX)'$. Is $X = x$ dan is $dy = \frac{\partial y}{\partial x} dx$; $y = a'x$ geeft $\frac{\partial y}{\partial x} = a'$. Ook geldt omgekeerd: $dy = sp(AdX) \Rightarrow A = \frac{\partial y}{\partial X}$. Analoog geldt:

B2. $$d^2y = d(dy) = d(sp\, \frac{\partial y}{\partial X}\, dX) = sp(d\, \frac{\partial y}{\partial X}\, dX) \;.$$

Is $X = x$ dan is $d^2y = dx'Hdx$ met $H := \{\frac{\partial^2 y}{\partial x_i \partial x_j}\}$ de matrix van Hesse ($H = H'$ als de afgeleiden continu zijn). Voldoende voor een maximum is $\frac{\partial y}{\partial X} = \hat{0}$ en $H$ negatief definiet.

B3. $$\frac{\partial |X|}{\partial X} = |X|X^{-1} \qquad \text{(X regulier ondersteld)} \;.$$

Bewijs: $|X| = \sum_i x_{ij}|X|_{ij}$, dus $\frac{\partial |X|}{\partial x_{ij}} = |X|_{ij}$. Nu is $(X^{-1})_{ij} = |X|_{ji}/|X|$ dus $|X|_{ji} = |X|(X^{-1})_{ij} = (\frac{\partial |X|}{\partial X})_{ij}$.

B4. $z = z(y(X))$ geeft $\frac{\partial z}{\partial X} = \frac{dz}{dy} \cdot \frac{\partial y}{\partial X}$ .

Een voorbeeld:

B5. $$\frac{\partial \log|X|}{\partial X} = \frac{1}{|X|}\,|X|X^{-1} = X^{-1} \,. \qquad (|X| > 0 \text{ ondersteld}).$$

B6. Door uitschrijven ziet men eenvoudig in dat

$$d(X+Z) = dX + dZ \;;\quad d(XZ) = (dX)Z + XdZ \;;$$

$$dX' = (dX)' = dX \text{ indien } X \text{ symmetrisch};\quad d \text{ sp } X = \text{sp } dX \,.$$

B7. $$dX^{-1} = -X^{-1}(dX)X^{-1} \,.$$

Bewijs:

$$0 = dI = d(XX^{-1}) = (dX)X^{-1} + Xd(X^{-1}) \,.$$

B8. Stel $y = x'Ax$ met $A$ symmetrisch, dan is $\frac{\partial y}{\partial x'} = 2Ax$, nl.

$$dy = dx'Ax + x'Adx = dx'Ax + dx'A'x = 2dx'Ax = \text{sp}(2Axdx') \,.$$

c. Toepassingen in de multivariate analyse.

Zij $\underline{x} \sim N_p(\mu,\Sigma^2)$ en $X_{n\times p}$ een steekproef van $\underline{x}$ ter grootte n.
Gevraagd maximum likelihoodschatters voor $\mu$ en $\Sigma^2$ (> 0 ondersteld).

Oplossing: De likelihoodfunktie (simultane kansdichtheid van $\underline{X}$) is volgens 3.1

$$L = |\Sigma^2|^{-\frac{1}{2}n}(2\pi)^{-\frac{1}{2}np} \exp[-\tfrac{1}{2}\sum_1^n (x_{i*}-\mu)'\Sigma^{-2}(x_{i*}-\mu)]$$

ofwel met $A := \Sigma^{-2}$

$$y = \log L = \text{const} + \tfrac{1}{2}n\log|A| - \tfrac{1}{2}\sum_1^n (x_{i*}-\mu)'A(x_{i*}-\mu) \,.$$

(1) schatter voor $\mu$:

$$\frac{\partial \log L}{\partial \mu'} = +\tfrac{1}{2}\sum_1^n 2A(x_{i*}-\mu) = A\sum_1^n (x_{i*}-\mu) = 0 \qquad \text{(zie B8)}$$

en daar $A > 0$, is

$$\underline{\hat{\mu}} = \sum_1^n \underline{x}_{i*}/n =: \underline{\bar{x}} \ .$$

Dit geeft inderdaad een maximum, nl.

$$dy = sp\{(A\,\Sigma\, x_{i*} - nA\mu)d\mu'\}$$

$$d^2y = sp(-\,nAd\mu d\mu') = -\,n\ sp(d\mu' A d\mu) < 0 \ .$$

(2) schatter voor $\Sigma^2$:

Nu wordt

$$y = c + \tfrac{1}{2}n\log|A| - \tfrac{1}{2}\sum_1^n (x_{i*} - \bar{x})' A (x_{i*} - \bar{x}) =$$

$$= c + \tfrac{1}{2}n\log|A| - \tfrac{1}{2}\ sp\ \sum (x_{i*} - \bar{x})(x_{i*} - \bar{x})' A$$

oftewel met

$$S^2 := \sum_1^n (x_{i*} - \bar{x})(x_{i*} - \bar{x})'/n$$

wordt dit

B9. $$y = \log L = c + \tfrac{1}{2}n\log|A| - \tfrac{1}{2}n\ sp(S^2A) \ .$$

Dit geeft: $dy = sp(\frac{\partial y}{\partial A}\,dA)$ oftewel

B10. $$dy = \tfrac{1}{2}n\ sp(A^{-1}dA - S^2dA) = \tfrac{1}{2}n\ sp[(A^{-1} - S^2)dA]$$

(A symmetrisch; B1, B5, B6).

B11. $$d^2y = -\,\tfrac{1}{2}n\ sp(A^{-1}dAA^{-1}dA) \ .$$

Het maximum van L (van logL) wordt bereikt voor $\underline{\hat{A}}^{-1} = \underline{S}^2 = \underline{\hat{\Sigma}}^2$.
Dit geeft inderdaad een maximum, daar $d^2y < 0$ volgens 1.68 en 1.69.
Het maximum wordt dan:

B12. $$L_{max} = |\hat{\Sigma}^2|^{-\frac{1}{2}n}(2\pi)^{-\frac{1}{2}np}\exp(-\tfrac{1}{2}np) \ .$$

Samengevat: de *maximum likelihoodschatters* voor $\mu$ en $\Sigma^2$ zijn resp.:

B13. $\underline{\hat{\mu}} = \underline{\bar{x}} := \sum_1^n \underline{x}_{i*}/n$ en $\underline{\hat{\Sigma}}^2 = \underline{S}^2 := \sum_1^n (\underline{x}_{i*} - \underline{\bar{x}})(\underline{x}_{i*} - \underline{\bar{x}})'/n$ .

In de faktoranalyse moet dezelfde funktie y uit B9 gemaximaliseerd worden, doch nu naar F en $V^2$. Daar is nl. $A^{-1} := \Sigma^2(\underline{x}) = F^2 + V^2$; $V^2$ is een diagonaalmatrix; $F^2 = F'F$.

Nu is $dA = d\Sigma^{-2} = -\Sigma^{-2}d\Sigma^2\Sigma^{-2}$ (zie B7).

En $d\Sigma^2 = dF^2 + dV^2 = (dF')F + F'dF + dV^2$

B10 wordt nu:

$$dy = \tfrac{1}{2}n\ \mathrm{sp}[(\Sigma^2 - S^2)(-\Sigma^{-2}d\Sigma^2\Sigma^{-2})] =$$
$$= \tfrac{1}{2}n\ \mathrm{sp}[(\Sigma^{-2}S^2\Sigma^{-2} - \Sigma^{-2})d\Sigma^2] =$$
$$= \tfrac{1}{2}n\ \mathrm{sp}[(\Sigma^{-2}S^2\Sigma^{-2}-\Sigma^{-2})dF^2 + (\Sigma^{-2}S^2\Sigma^{-2} - \Sigma^{-2})dV^2].$$

Nu is $\mathrm{sp}[(\Sigma^{-2}S^2\Sigma^{-2}-\Sigma^{-2})(dF'F + F'dF)] =$

$= \mathrm{sp}[2(F\Sigma^{-2}S^2\Sigma^{-2} - F\Sigma^{-2})dF']$ oftewel

$$dy = \tfrac{1}{2}n\ \mathrm{sp}[2(F\Sigma^{-2}S^2\Sigma^{-2}-F\Sigma^{-2})dF' + (\Sigma^{-2}S^2\Sigma^{-2} - \Sigma^{-2})dV^2].$$

Dus maximaal voor:

B14. $\hat{F}\hat{\Sigma}^{-2}S^2\hat{\Sigma}^{-2} - \hat{F}\hat{\Sigma}^{-2} = 0$ geeft $\hat{F}\hat{\Sigma}^{-2}S^2 = \hat{F}$.

B15. $(\hat{\Sigma}^{-2}S^2\hat{\Sigma}^{-2} - \hat{\Sigma}^{-2})_{ii} = o \quad i = 1,\ldots,p$ dus alleen de diagonaal, omdat $V^2$ diagonaal is.

Appendix D: De gegeneraliseerde inverse

a. Definitie: Zij $A \in \mathfrak{M}^r_{p,q}$. Een gegeneraliseerde inverse van A is een matrix $A^-$ van de orde $q \times p$ zodat $AA^-A = A$.

Opmerkingen: Er zijn, zoals nog zal blijken, bij een gegeven A oneindig veel gegeneraliseerde inversen. Alleen als A regulier is geldt $A^- = A^{-1}$. Men kan $A^-$ uniek maken door nog andere voorwaarden op te leggen. Zo ontstaat de Moore-Penrose- of ook wel pseudo-inverse $A^+$. (Dan geldt o.a. $A^+AA^+ = A^+$.)

D1. $r(A^-A) = r(AA^-) = r(A) =: r$ (zie 1.47).
Evident is dat $r(A^-) \geq r(A)$ (met (1.44).
$A^-A$ en $AA^-$ evenals $I_q - A^-A$ zijn idempotent (ga na!).
Is $r = q$ dan is bv. een $A^- = (A'A)^{-1}A'$ en $A^-A = I_q$ (zie 1.80).
Is $r = p$ dan is bv. een $A^- = A'(AA')^{-1}$ en $AA^- = I_p$.

Notatie: $I^o := A^-A$. Dus $I^{o2} = I^o$ en $AI^o = A$.

D2. $N(A) = \langle I - I^o \rangle$ nl: $A(I - I^o) = A - AI^o = 0$ dus $\langle I - I^o \rangle \subset N(A)$
$\dim N(A) = q - r$ en $r(I - I^o) = sp(I - I^o) = q - r$ (zie 1.94; 1.39 en D1).

b. Existentie 1:

D3. Zij weer $A \in \mathfrak{M}^r_{p,q}$. In 1.81 is $A = W\Lambda^{\frac{1}{2}}V'$. Definieer $A^- := V\Lambda^{-\frac{1}{2}}W'$, dan blijkt inderdaad $AA^-A = A$. In feite is dit de $A^+$.
We zullen echter vaak een andere gegeneraliseerde inverse beschouwen.

D4. Existentie 2:

We permuteren de rijen en kolommen van A zò, dat de leidende ondermatrix van de ontstane $\tilde{A}$ regulier is (zie 1.38 en 1.46): $\tilde{A} = \tilde{I}_1 A \tilde{I}_2$ oftewel $A = \tilde{I}_1' \tilde{A} \tilde{I}_2'$; $r(\tilde{A}) = r(A)$ (met 1.44). Indien er een $\tilde{A}^-$ bestaat, is er een $A^- := \tilde{I}_2 \tilde{A}^- \tilde{I}_1$ (ga na met 1.38).
Nu stel

$$\tilde{A} = \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{pmatrix},$$

Opmerking: is $r(\tilde{A}_{11}) = r(\tilde{A})$, dan bestaat $\tilde{A}_{12}$ niet. Doch dan zijn we al klaar nl. zie D1 met $r = q$.

dan is

$$\tilde{A}^{-} := \begin{pmatrix} \tilde{A}_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}$$

een gegeneraliseerde inverse van $\tilde{A}$ (ga na met 1.36).

Elke matrix heeft dus een gegeneraliseerde inverse!

In het volgende zullen we steeds, zonder verlies van algemeenheid, onderstellen dat de leidende $r \times r$ ondermatrix $A_{11}$ regulier is.

c. Samenhang verschillende gegeneraliseerde inversen

Zij steeds $A \in \mathfrak{M}^{r}_{p,q}$ en $r(A_{11}) = r$;

$$A^{-} = G = \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix}$$

met $G_{11}$ van de orde $r \times r$.

Nu is $A = LA_{11}M$ met

$$L := \begin{pmatrix} I_r \\ A_{21}A_{11}^{-1} \end{pmatrix}$$

en $M := (I_r \quad A_{11}^{-1}A_{12})$.

Ondersteld is $r < \min(p,q)$. Is echter $r = q$, dan bestaat $A_{12}$ niet en kan men een analoge afleiding geven, die resulteert in

$$B = G_{12}A_{21}A_{11}^{-1}$$

(dus alsof $A_{12} = 0$).

$AGA = A$ wordt $LA_{11}MGLA_{11}M = LA_{11}M$ oftewel met 1.60ii), daar L volle kolom-, M volle rijrang heeft: $MGL = A_{11}^{-1}$.

M, G en L hierin gesubstitueerd in gepartitioneerde vorm geeft:

D5. $$G_{11} + A_{11}^{-1}A_{12}G_{21} + G_{12}A_{21}A_{11}^{-1} + A_{11}^{-1}A_{12}G_{22}A_{21}A_{11}^{-1} = A_{11}^{-1}.$$

We hebben hier een noodzakelijke voorwaarde waaraan de ondermatrices van G moeten voldoen. Het is ook een voldoende voorwaarde; elke G die aan D5 voldoet is een $A^{-}$, daar elke stap terug geldt en tenslotte weer $AGA = A$ oplevert.

$G_{11}$ oplossen uit D5 geeft: $G_{11} = A_{11}^{-1} - B$ met

$$B := A_{11}^{-1}A_{12}G_{21} + G_{12}A_{21}A_{11}^{-1} + A_{11}^{-1}A_{12}G_{22}A_{21}A_{11}^{-1} .$$

Dan is elke

$$G_{q\times p} := \begin{pmatrix} A_{11}^{-1} - B & G_{12} \\ G_{21} & G_{22} \end{pmatrix}$$

met $G_{12}$, $G_{21}$, $G_{22}$ willekeurig (wel van de juiste orde) een $A^-$.

Opmerkingen:

i) Voor $G_{12} = 0$, $G_{21} = 0$, $G_{22} = 0$ ontstaat

$$A^- = \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} .$$

D6. ii) Elke A heeft een $A^-$ van volle rang (als p = q dus een reguliere $A^-$) nl. stel $p \geq q > r$. Neem $G_{21} = 0$, $G_{12} = -A_{11}^{-1}A_{12}G_{22}$ (dan is B = 0 en dus $G_{11} = A_{11}^{-1}$) en neem $G_{22}$ met volle rijrang $q - r$.
Dan is $r(G_{11}) = r(A_{11}) = r$ en $r(G) = r + (q - r) = q$.

Er bestaat een eenvoudige relatie tussen de verschillende gegeneraliseerde inversen van een matrix A. Met $A^-$ is nl. ook een gegeneraliseerde inverse van A:

D7. $$G := A^-AA^- + (I_q - A^-A)X + Y(I_p - AA^-) \text{ (ga na)} .$$

Hierin zijn X en Y willekeurig van de orde $q \times p$.
Bij gegeven $A^-$ geeft D7 ook alle gegeneraliseerde inversen van A.
Zij bv. $G^*$ een gegeneraliseerde inverse van A, dan krijgt men deze uit D7 door in te vullen $X = G^*$ en $Y = A^-AG^*$ (ga na).
Met de gebruikelijke partitionering, nu ook voor X en Y, en met

$$A^- := \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}$$

wordt D7:

D8. $$G = \begin{pmatrix} A_{11}^{-1} - A_{11}^{-1}A_{12}X_{21} - Y_{12}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}X_{22} + Y_{12} \\ X_{21} - Y_{22}A_{21}A_{11}^{-1} & X_{22} + Y_{22} \end{pmatrix}.$$

Evenals D7, geeft D8 alle gegeneraliseerde inversen van A.
Dat

$$G = \begin{pmatrix} A_{11} - B & G_{12} \\ G_{21} & G_{22} \end{pmatrix}$$

ook alle gegeneraliseerde inversen van A oplevert, is nu in te zien door $G_{12}$, $G_{21}$ en $G_{22}$ te nemen als in D8, dan wordt $A_{11} - B$ juist de $G_{11}$ uit D8. Zoals blijkt uit D8, wordt de willekeur in G bepaald door een gedeelte van de matrices X en Y; blijkbaar spelen $X_{11}$, $X_{12}$, $Y_{11}$, $Y_{21}$ geen rol.

## d. Gegeneraliseerde inversen bij semi positief definiete matrices

D9. Stelling: elke symmetrische A heeft een symmetrische $A^-$.

Bewijs: $AA^-A = A$ transponeren geeft $A(A^-)'A = A$, dus $(A^-)'$ is ook gegeneraliseerde inverse. Definieer $B := \frac{1}{2}[A^- + (A^-)']$, deze is symmetrisch en $ABA = A$.

D10. Zij $G = (X'X)^-$. Dan geldt: $XGX' =: P_X$ is de orthogonale projector op <X> en dus onafhankelijk G.

Bewijs:

i) Met D9 $X'XG'X'X = X'X$ dus (zie 1.60) $X'XG'X' = X'$. Transponeren geeft $XGX'X = X$ oftewel $P_XX = X$. Bovendien is $r(P_X) = r(X)$. Ga na.

ii) $P_XP_X = (XGX')(XGX') = (XGX'X)GX' = XGX' = P_X$ dus idempotent.

iii) Stel F is ook $(X'X)^-$; $X = XGX'X = XFX'X$ en met 1.60 volgt $XGX' = XFX'$ oftewel invariant met betrekking tot G.

iv) $P_X' = (XGX')' = XG'X' = XGX' = P_X$ dus $P_X$ symmetrisch.
Tezamen geeft dit, dat $P_X$ een orthogonale projector is op <X> en (dus) onafhankelijk G.

We zullen nu, naast de vorige, enkele methoden bespreken voor de constructie van een gegeneraliseerde inverse van X'X.

D11. <u>Stelling</u>: Zij $X \in \mathcal{M}_{np}^{p-m}$ en

$$S := \begin{pmatrix} X'X & B' \\ B & 0 \end{pmatrix}$$

regulier, $r(S) = p + m$ (zie 1.55).
Is

$$S^{-1} =: \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$$

dan zijn $C_{11}$ en $(X'X + B'B)^{-1}$ gegeneraliseerde inversen van $X'X$.

<u>Bewijs</u>:

$$SS^{-1} = \begin{pmatrix} I_p & 0 \\ 0 & I_m \end{pmatrix}$$

geeft $X'XC_{11} + B'C_{21} = I_p$ en $BC_{11} = 0$.
$D'X'XC_{11} + D'B'C_{21} = D'$ met $D$ uit 1.42 zodat $D'X'X = 0$ en $D'B'$ regulier.
Oftewel $C_{21} = (D'B')^{-1}D'$. Verder is $X'XC_{11} = I - B'(D'B')^{-1}D'$ en navermenigvuldiging met $X'X$ geeft: $X'XC_{11}X'X = X'X - B'(D'B')^{-1}D'X'X = X'X$, dus $C_{11} = (X'X)^-$.

Nu is $(X'X + B'B)[C_{11} + D(D'B'BD)^{-1}D'] = I$ (ga na door uit te werken!).
Hieruit volgt $(X'X + B'B)^{-1} = C_{11} + D(D'B'BD)^{-1}D'$ oftewel
$X'X(X'X + B'B)^{-1}X'X = X'XC_{11}X'X + 0 = X'X$ dus eveneens $(X'X + B'B)^{-1}$ is een $(X'X)^-$.

<u>Opmerking</u>: $X'X + B'B = \binom{X}{B}^{\underline{2}} > 0$ oftewel $X'X$ heeft een $(X'X)^- > 0$.

e. <u>Consistent stelsel lineaire vergelijkingen</u>

D12. <u>Stelling</u>: $Ax = y$ consistent; $x = Gy$ is een oplossing $\Longleftrightarrow G = A^-$.
Hier is bedoeld: $x = Gy$ is een oplossing voor <u>alle</u> toelaatbare $y$, dus voor alle $y \in \langle A \rangle$.

Bewijs:

$\Leftarrow$: $Ax = y$ heeft oplossing $x^*$, dus $Ax^* = y$; $AGA = A$, dus $AGAx^* = Ax^*$, of $AGy = y$, dus $Gy$ is oplossing bij $y$.

$\Rightarrow$: consistent, d.w.z. $y \in \langle A \rangle$. Neem $y = a_{*j}$. Gegeven is $Gy = Ga_{*j}$ is een oplossing van $Ax = a_{*j}$. Dus $AGa_{*j} = a_{*j}$ (voor elke j) dus ook $AG(a_{*1} \ldots a_{*q}) = (a_{*1} \ldots a_{*q})$ of $AGA = A$ dus $G = A^-$.

D13. Stelling: Zij $A \in \mathfrak{M}^r_{p,q}$ en $Ax = y$ consistent.
Alle oplossingen zijn $x^o = A^-y + (I_q - I^o)z$, z willekeurig.

Bewijs:

1) $Ax^o = AA^-y + (A - AA^-A)z = y$ (D12).
2) Stel $x^*$ is een oplossing, neem $z = x^*$, $x^o = A^-y + (I - A^-A)x^* = x^*$.

Opmerking: één $A^-$ genereert dus alle oplossingen.

D14. Stelling: Als G de verzameling van alle gegeneraliseerde inversen van A doorloopt, dan genereert $x^o = Gy$ alle oplossingen van het consistente stelsel $Ax = y$ ($\neq 0$).

Bewijs: Zij $A^-$ een gegeneraliseerde inverse van A, dan geeft $x^o = A^-y + (I - I^o)z$ alle oplossingen (D13). Neem $z = Xy$ (zie 1.43) dan wordt dit

$$x^o = A^-y + (I - I^o)Xy = [A^- + (I - I^o)X]y =: Gy \ .$$

G is nl. een gegeneraliseerde inverse van A (ga na!).

D15. De volgende uitspraken zijn equivalent (zie vraagstuk D 18).

a) $AGA = A$, $(GA)' = GA$
b) $AGA = A$, $(GA)'(I - GA) = 0$
c) $AGA = A$, $\langle GA \rangle \subset \langle A' \rangle$ (er geldt het =)
d) $GAA' = A'$
e) $GA = P_A$, (er geldt zelfs dat deze orthogonaal is).

D16. <u>Definitie</u>: Zij $Ax = y$ consistent.

$x_m$ heet een <u>minimum</u> <u>norm</u> oplossing als $\begin{cases} Ax_m = y \text{ en} \\ \| x_m \| = \min\limits_{Ax=y} \| x \| \end{cases}$

Hierin is $\| x \|^2 = x'x$.

$x = A^- y$; noteer $x_m = A^-_m y$.

D17. <u>Definitie</u>: Zij $Ax = y$ consistent.

$x_{m(v)}$ heet een minimum V-norm oplossing als $\begin{cases} Ax_{m(v)} = y \text{ en} \\ \| x_{m(v)} \|_v = \min\limits_{Ax=y} \| x \|_v \end{cases}$

Hierin is $\| x \|_v = x'Vx$ en $V \geq 0$

$x = A^- y$; noteer $x_{m(v)} = A^-_{m(v)} y$.

D18. <u>Stelling</u>:

$$G = A^-_{m(v)} \Leftrightarrow [AGA = A \wedge (VGA)' = VGA]; \text{ in het bijzonder voor } V = I$$

$$G = A^-_m \Leftrightarrow [AGA = A \wedge (GA)' = GA]$$

<u>Bewijs</u>: Voor een oplossing van $Ax = y$, bv. $x = Gy$, is noodzakelijk en voldoende $AGA = A$ (zie D 12).

Alle oplossingen zijn: $x = Gy + (I - GA)z$ met $z$ willekeurig (zie D 13).

$y \in \langle A \rangle$ dus $y = Ab$. Stel $V = \Sigma'\Sigma$.

Nu moet $\| Gy \|_v \leq \| Gy + (I - GA)z \|_v$ oftewel door $\| x \|^2_v = \| \Sigma x \|^2$

$$\| \Sigma GAb \| \leq \| \Sigma GAb + \Sigma(I - GA)z \| \quad \text{voor alle } b, z.$$

Dus $(\Sigma GA)'(\Sigma - \Sigma GA) = 0$; $(\Sigma GA)'\Sigma = (\Sigma GA)'\Sigma GA$ (symmetrisch)

dus $(\Sigma GA)'\Sigma = \Sigma'\Sigma GA$, dus $A'G'\Sigma'\Sigma = \Sigma'\Sigma GA$, $(VGA)' = VGA$. □

D19. $x_m$ is uniek nl. $Ax = y$ heeft als oplossingen een lineaire nevenruimte.
$\| x \|$ minimaal d.w.z. projecteer 0 op deze ruimte, dat geeft een éénduidige oplossing.

Is $V > 0$, dan is $x_{m(v)}$ ook uniek. Stel $V = \Sigma'\Sigma$ en $\Sigma x = z$.

$Ax = y$ wordt: $A\Sigma^{-1}z = y$ of $Bz = y$ met $\|x\|_V = \|z\|$ minimaal, dus vorig probleem.

$A_m^-$ is niet uniek, o.a. voldoet $A_m^- := A'(AA')^-$. Dit volgt uit D 15 a ~ d.

D20. <u>Hulpstelling</u>:

$$\begin{pmatrix} V & A' \\ A & 0 \end{pmatrix}^- = \begin{pmatrix} C_1 & C_2 \\ C_3 & -C_4 \end{pmatrix} \text{ dan is } C_2 = A^-_{m(v)}$$

<u>Bewijs</u>: We zoeken de minimum V-norm oplossing van $Ax = y$ met de Lagrange-methode:

$$f(x) := x'Vx + 2\lambda'(Ax - y); \quad 2Vx + 2A'\lambda = 0$$

$$\frac{\delta f}{\partial x'} = 0 \qquad\qquad Ax - y = 0$$

oftewel anders geschreven

$$\begin{pmatrix} V & A' \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}; \quad \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} V & A' \\ A & 0 \end{pmatrix}^- \begin{pmatrix} 0 \\ y \end{pmatrix} = \begin{pmatrix} C_2 & y \\ -C_4 & y \end{pmatrix} \quad \text{dus } x = C_2 y = A^-_{m(v)} y$$

□

Gevolg met D 18: $AC_2A = A$ èn $(VC_2A)' = VC_2A$.

D21. <u>Theorema</u>:

Zij

$$\begin{pmatrix} V & X \\ X' & 0 \end{pmatrix}^- = \begin{pmatrix} C_1 & C_2 \\ C_3 & -C_4 \end{pmatrix}; \; V_{n\times n} \geq 0; \; X_{n\times m};$$

Er geldt:

a) $XC_3X = X$ en $(VC_3'X') = XC_3V$ m.a.w. $C_3' = (X')^-_{m(v)}$. Analoog
$XC_2'X = X$ en $(VC_2X') = XC_2'V$ $\qquad C_2 = (X')^-_{m(v)}$

b) $X'C_1X = VC_1X = X'C_1V = 0$ (dus ook $VC_1'X = X'C_1'X = 0$)

c) $VC_2X' = VC_3'X' = XC_4X'$ (dus laatste volgens a ōōk symmetrisch)

d) $VC_1VC_1V = VC_1V$;

e) $\begin{pmatrix} C_1 \\ C_3 \end{pmatrix}$ is een geg. inverse van $(V \,\vdots\, X)$; $sp(VC_1) = r(V \,\vdots\, X) - r(X)$

f) $\begin{pmatrix} C_1 & C_3' \\ C_2' & -C_4 \end{pmatrix}$ is ook een geg. inverse van $\begin{pmatrix} V & X \\ X' & 0 \end{pmatrix}$

g) $VC_1V$ is symmetrisch

<u>Bewijs</u>:

a) Dit volgt direct uit D20 met $A = X'$, dus $X'C_2X' = X'$ en $(VC_2X')' = VC_2X'$

Daar $\begin{pmatrix} V & X \\ X' & 0 \end{pmatrix}$ symmetrisch is, is $\begin{pmatrix} C_1' & C_3' \\ C_2' & -C_4' \end{pmatrix}$ ook een geg. inverse, dus geldt

alles ook voor $C_3'$, m.a.w. $C_2 \leftrightarrow C_3'$.

b) $\begin{pmatrix} V & X \\ X' & 0 \end{pmatrix} \begin{pmatrix} C_1 & C_2 \\ C_3 & -C_4 \end{pmatrix} \begin{pmatrix} V & X \\ X' & 0 \end{pmatrix} = \begin{pmatrix} V & X \\ X' & 0 \end{pmatrix}$

geeft de relaties:

1) $VC_1V + VC_2X' + XC_3V - XC_4X' = V$

2) $VC_1X + XC_3X = X$

3) $X'C_1V + X'C_2X' = X'$

4) $X'C_1X = 0$

Uit 4) volgt dus $X'C_1X = 0$; Uit 2) met a volgt $VC_1X = 0$
Uit 3) met a volgt $X'C_1V = 0$.

c) 1) achtervermenigv. met $C_2X'$ geeft:

$$VC_1\underbrace{VC_2X'} + VC_2\underbrace{X'C_2X'} + XC_3\underbrace{VC_2X'} - XC_4\underbrace{X'C_2X'} = VC_2X'$$

$$\underbrace{VC_1X}_{\substack{\| \\ 0}}C_2'V + VC_2X' + \underbrace{XC_3X}_{\substack{\| \\ X}}C_2'V - XC_4X' = VC_2X'$$

Dus $XC_2'V = XC_4X'$ dus met a $VC_2X' = XC_4X'$. Analoog

$$\underbrace{XC_3}VC_2X' = XC_4X'$$

$$VC_3'\underbrace{X'C_2}X' = VC_3'X' = XC_4X' \qquad \square$$

d) 1) Achtermenigv. met $C_1V$ geeft:

$$VC_1VC_1V + VC_2\underbrace{X'C_1}_{0}V + \underbrace{XC_3}VC_1V - XC_4\underbrace{X'C_1}_{0}V = VC_1V$$

$$VC_3\underbrace{X'C_1}V = 0$$

dus $VC_1VC_1V = VC_1V$ $\square$

e) met c wordt 1): $VC_1V + XC_3V = V$

$$(V \vdots X)\begin{pmatrix} C_1 \\ C_3 \end{pmatrix}(V \vdots X) = (\underbrace{VC_1V + XC_3V}_{V} \vdots \underbrace{VC_1X}_{0} + \underbrace{XC_3X}_{X}) = (V \vdots X) \text{ q.e.d.}$$

$AA^-$ is idempotent dus rang = spoor, dus $r\{(V \vdots X)\begin{pmatrix} C_1 \\ C_3 \end{pmatrix}\} = sp[VC_1 + XC_3] = r(V \vdots X)$

daar $r(AA^-) = r(A)$ zie D1.

$sp[VC_1 + XC_3] = sp[VC_1) + sp(XC_3) = sp(VC_1) + r(X)$ daar $sp(XC_3) = r(XC_3)$

$= r(X)$ (uit a)

oftewel $sp(VC_1) = r(V \quad X) - r(X)$ $\square$

f) Volgt direct door uit te werken.

g) $\begin{pmatrix} C_1' & C_3' \\ C_2' & -C_4' \end{pmatrix}$ is ook een geg. inverse, dus in dat geval zou 1^e regel onder

bewijs van e) zijn: $VC_1'V + XC_2'V = V$.

Nu is $V$, $XC_2'V$ symmetrisch, dus ook $VC_1'V = VC_1V$. $\square$

Appendix D: <u>De gegeneraliseerde inverse</u>.

D.1. Is A regulier dan is $A^- = A^{-1}$. Bewijs.

D.2. Bewijs $r(A^-) \geq r(A)$. Wanneer geldt zeker het gelijkteken?

D.3. Bewijs: $A^- := \tilde{I}_2 \widetilde{A^-} \tilde{I}_1$ is een gegeneraliseerde inverse van A; zie (D.4).

D.4. Bepaal een gegeneraliseerde inverse van $A = \begin{pmatrix} 31 \\ 62 \end{pmatrix}$. Tevens een <u>reguliere</u> gegeneraliseerde inverse.
Ook alle gegeneraliseerde inversen in de vorm van één matrix geschreven.

D.5. $A = \begin{pmatrix} 21 \\ 11 \\ 01 \end{pmatrix}$ Bepaal een $A^-$. Tevens $(A'A)^{-1}A'$-

D.6. Bepaal alle gegeneraliseerde inversen van $A = \begin{pmatrix} 10 \\ 01 \\ 11 \end{pmatrix}$ in één matrix geschreven.

D.7. $A = \begin{pmatrix} 3 & 2 & \frac{1}{2} \\ 6 & 4 & 1 \\ 0 & 8 & 2 \end{pmatrix}$ Bepaal een $A^-$.

D.8. Voor reguliere matrices geldt $(AB)^{-1} = B^{-1}A^{-1}$.
Bewijs: $(AB)^{-1} = B^-A^- \Leftrightarrow A^-ABB^-$ idempotent.
Dus in het bijzonder als B regulier is geldt $(AB)^- = B^{-1}A^-$ .

D.9. Zij $A \in M^p_{pq}$ en $B \in M^p_{pp}$. Bewijs: $[A(A'BA)^-A']^{-1} = B$.

D.10. Zij $A_{pq}$ heeft een gegeneraliseerde inverse zodat $AA^- = I$.
Welke voorwaarde is noodzakelijk en voldoende voor A opdat er zo'n $A^-$ is?

D.11. Zij $A \in M^p_{pq}$. Bewijs $A(A'A)^-A' = I$.
Algemeen geldt: $A(A'A)^-A'A = A$ en $A(A'A)^-A' = P_A$, zie (D.10).

D.12. Stel $A_{pq}$ met $p \neq q$. Bewijs: er is géén gegeneraliseerde inverse $A^-$ zodat $AA^- = I_p$ èn $A^-A = I_q$. (is er wel zo'n $A^-$ dan is p = q en dus A regulier).

D.13. Bewijs: $r[X(X'X)^-X'] = r(X)$.

D.14. $(A^-)' = (A')^-$. Bewijs.

D.15. $x \in \langle A\rangle \Leftrightarrow AA^-x = x$. Gevolg: $Ax = y$ is consistent d.e.s.d. als $AA^-y = y$.

D.16. $\{A^-\} = \{B^-\} \Leftrightarrow A = B$ d.w.z. A en B zijn gelijk indien hun klassen van gegeneraliseerde inversen gelijk zijn.

D.17. $BA^-A = B \Leftrightarrow \exists D$ zodat $B = DA$.

D.18. Bewijs (D.15) uit diktaat.

Appendix E: Het Kronecker- of directproduct van matrices

a. Definitie: $A \otimes B := \{a_{ij}B\}$. Is $A_{p\times q}$, $B_{r\times s}$ dan is $A \otimes B$ een matrix van de orde $pr \times qs$.

Elk element $a_{ij}$ van A wordt met de gehele B vermenigvuldigd.

Opmerking: In het algemeen is $A \otimes B \neq B \otimes A$.

$A \otimes I = \{a_{ij}I\}$ maar

$$I \otimes A = \begin{pmatrix} A & & 0 \\ & \ddots & \\ 0 & & A \end{pmatrix} .$$

$I_p \otimes I_q = I_q \otimes I_p = I_{pq}$.

Voor vectoren geldt: $a' \otimes b = b \otimes a' = ba'$.

b. Eigenschappen:

E1. Partitioneren: $(A_1 A_2 A_3) \otimes B = (A_1 \otimes B,\ A_2 \otimes B,\ A_3 \otimes B)$.

E2. $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ indien de producten AC en BD bestaan, nl.:

$(A \otimes B)(C \otimes D) = \{a_{ik}B\}\{c_{kj}D\} = \{a'_{i*}c_{*j}BD\} = (AC) \otimes (BD)$.

Gevolg 2a: A, B idempotent $\Rightarrow$ $A \otimes B$ idempotent.

E3. $A \otimes (B + C) = A \otimes B + A \otimes C$. Deze volgen direct uit de definitie.

$(B + C) \otimes A = B \otimes A + C \otimes A$.

$(A \otimes B) \otimes C = A \otimes (B \otimes C)$ (associatief).

E4. $(A \otimes B)' = A' \otimes B'$, nl. $\{a_{ij}B\}' = \{a_{ji}B'\} = A' \otimes B'$.

Gevolg 4a: A, B symmetrisch $\Leftrightarrow$ $A \otimes B$ symmetrisch.

4b: A, B orthogonale projecties $\Rightarrow$ $A \otimes B$ orthogonale projectie (met 2a en 4a).

4c: A, B orthogonaal $\Rightarrow$ $A \otimes B$ orthogonaal.

E5. Zij $A_{p\times p}$ met eigenwaarden $\lambda_i$ en eigenvectoren $x_i$, $i = 1,\dots,p$.
En $B_{q\times q}$ met eigenwaarden $\mu_j$ en eigenvectoren $y_j$, $j = 1,\dots,q$.

Dan heeft $A \otimes B$ als eigenwaarden $\lambda_i\mu_j$ met eigenvectoren $x_i \otimes y_j$, nl.:

$$(A \otimes B)(x_i \otimes y_j) = Ax_i \otimes By_j = \lambda_i\mu_j(x_i \otimes y_j) .$$

E6. $|A_{p\times p} \otimes B_{q\times q}| = \prod_{ij} \lambda_i\mu_j = \prod_i \lambda_i^q \prod_j \mu_j^p = |A|^q|B|^p$ (pas op, p en q wisselen!).

E7. $\mathrm{sp}(A_{p\times p} \otimes B_{q\times q}) = \sum_{ij} \lambda_i\mu_j = \sum_i \lambda_i \sum_j \mu_j = \mathrm{sp}\,A\ \mathrm{sp}\,B$.

E8. $r(A \otimes B) = r(A)r(B)$.

Bewijs: De grootste reguliere submatrix van A, stel $\bar{A}$, heeft de afmeting $r(A)$.
Analoog van B, stel $\bar{B}$, de afmeting $r(B)$. Nu is $\bar{A} \otimes \bar{B}$ submatrix van $A \otimes B$ en deze is regulier daar $|\bar{A} \otimes \bar{B}| = |\bar{A}|^{r(B)}|\bar{B}|^{r(A)} \neq 0$.
Elke grotere submatrix van $A \otimes B$ is singulier, nl. stel deze $A^*_{kk}$ en $B^*_{nn}$ dus $|A^* \otimes B^*| = |A^*|^n|B^*|^k = 0$ voor $k > r(B)$ of $n > r(A)$. De afmeting van $\bar{A} \otimes \bar{B}$ is juist $r(A)r(B)$.

Gevolg 8a: $A \otimes B$ regulier $\Longleftrightarrow$ A, B regulier (en dus vierkant!).
Bewijs $\Leftarrow$: Triviaal.
Bewijs $\Rightarrow$: Stel $A_{pq} \otimes B_{rs}$ regulier. Dan is dus $pr = qs = r(A)r(B)$. Stel $r(A) = p$ (< p uitgesloten anders $r(B) > r$). Dus $r(B) = r$. Als $q > p$ of $s > r$ is $pr \neq qs$, dus $p = q$ en $r = s$.
8b: $A,B > 0 \Rightarrow A \otimes B > 0$ (met 4a en 8a).

E9. $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$. Uit 8a volgt dat A en B regulier moeten zijn. Eveneens geldt voor een gegeneraliseerde inverse $(A \otimes B)^- = A^- \otimes B^-$.

c. De vec-notatie

Definitie: $\vec{A} := (a'_{*1} \ \ldots \ a'_{*q})'$, d.w.z. de kolommen van A achtereenvolgens onder elkaar genoteerd. Is A van de orde $p \times q$, dan is $\vec{A}$ (ook wel vec A genoteerd) een kolomvector met pq componenten. In het bijzonder is $\vec{a} = \vec{a}' = a$.

E10. Stelling: $\overrightarrow{A+B} = \vec{A} + \vec{B}$. Triviaal (ga na).

E11. Stelling: $\overrightarrow{ABC} = (C' \otimes A)\vec{B}$.

Bewijs: $(ABC)_{*j} = (Ab_{*1} \ \ldots \ Ab_{*s})c_{*j} = \sum_k c_{kj}Ab_{*k}$. En

$$(C' \otimes A)\vec{B} = \{c_{kj}A\}\begin{pmatrix} b_{*1} \\ \vdots \\ b_{*s} \end{pmatrix} = \begin{Bmatrix} \vdots \\ \sum_k c_{kj}Ab_{*k} \\ \vdots \end{Bmatrix} . \qquad \text{q.e.d.}$$

Gevolg 11a: $Y = AX$ geeft $\vec{Y} = (I \otimes A)\vec{X}$ en
$Y = XA$ geeft $\vec{Y} = (A' \otimes I)\vec{X}$.

E12. Stelling: $sp(A'B) = \sum_i a'_{*i}b_{*i} = (\vec{A},\vec{B})$ zodat $\|A\| = \|\vec{A}\|$ en $(A,B) = (\vec{A},\vec{B})$ (zie 1.8 en 1.9).

d. Enkele toepassingen

Definitie: VAR $\underline{X}$ := VAR $\underline{\vec{X}}$ (pas op: VAR $\underline{X} \neq$ VAR $\underline{X}'$).

E13. Stelling: $\underline{Y} = A\underline{X}$ geeft: VAR $\underline{Y} = (I \otimes A)$VAR $\underline{X}(I \otimes A')$
$\underline{Y} = \underline{X}A$ geeft: VAR $\underline{Y} = (A' \otimes I)$VAR $\underline{X}(A \otimes I)$, (met 2.37 en E11a).

E14. Voorbeeld 1: Zij $\underline{y} = X\beta + \underline{e}$ met $\underline{e} = \chi_n\sigma$ dus $\mathcal{E}\underline{e} = 0$ en var $\underline{e} = \sigma^2 I_n$. We weten nu:
$\underline{b} = (X'X)^{-1}X'\underline{y}$ en var $\underline{b} = \sigma^2(X'X)^{-1}$.
Beschouwen we nu simultaan q vergelijkingen, dan ontstaat:

$$\underline{Y} = XB + \underline{E} \text{ met } \underline{E} \simeq \chi_{n,q}\Sigma_{qq} \text{ dus } \mathcal{E}\underline{E} = 0 .$$

Nu blijkt:

$$\underline{\hat{B}} = (X'X)^{-1}X'\underline{Y} \text{ met VAR } \underline{\hat{B}} = \Sigma^2 \otimes (X'X)^{-1} \text{ en VAR } \underline{E} = \Sigma^2 \otimes I_n \,.$$

Bewijs:

i) VAR $\underline{E} = (\Sigma' \otimes I_n)(I_q \otimes I_n)(\Sigma \otimes I_n) = \Sigma^2 \otimes I_n$ (met E2, E13).

ii) $\underline{\vec{Y}} = (I \otimes X)\vec{B} + \underline{\vec{E}}$ dus $\underline{\hat{\vec{B}}} = [(I \otimes X')(I \otimes X)]^{-1}(I \otimes X')\underline{\vec{Y}} =$

$= [I \otimes (X'X)^{-1}][I \otimes X']\underline{\vec{Y}} = [I \otimes (X'X)^{-1}X']\underline{\vec{Y}}$ oftewel met E11

$\underline{\hat{B}} = (X'X)^{-1}X'\underline{Y}$.

iii) VAR $\underline{\hat{\vec{B}}} = [I \otimes (X'X)^{-1}X'][\Sigma^2 \otimes I_n][I \otimes X(X'X)^{-1}] = \Sigma^2 \otimes (X'X)^{-1}$.

Volgens E8b is $\Sigma^2 \otimes (X'X)^{-1} > 0$.

E15. Voorbeeld 2: Beschouw het volgende variantie-analyse model:

$$\underline{y}_{ij} = \mu + \alpha_i + \gamma_j + \underline{e}_{ij} \,, \quad i = 1,\ldots,p,\ j = 1,\ldots,q \,.$$

Dit is te schrijven als:

$$\underline{Y} = \mu U + \alpha u_q' + u_p \gamma' + \underline{E} \,.$$

Dus

$$\underline{\vec{Y}} = \mu u_{pq} + (u_q \otimes I_p)\alpha + (I_q \otimes u_p)\gamma + \underline{\vec{E}}$$

oftewel $\underline{\vec{Y}} = X\beta + \underline{\vec{E}}$ met $\beta := (\mu,\alpha',\gamma')'$ en $X := (u_{pq} u_q \otimes I_p I_q \otimes u_p)$.

E16. Voorbeeld 3: Zij $Ax = y$ consistent, dan worden alle oplossingen $x^o = A^- y + (I - A^- A)z$. Beschouw nu simultaan het consistente stelsel $AX = Y$. Wat zijn nu alle oplossingen?

$(I \otimes A)\vec{X} = \vec{Y}$ geeft $\vec{X}^o = (I \otimes A^-)\vec{Y} + [I - (I \otimes A^-)(I \otimes A)]\vec{Z} =$

$$= (I \otimes A^-)\vec{Y} + [I - (I \otimes A^- A)]\vec{Z} \,,$$

oftewel

$$X^o = A^- Y + (I - A^- A)Z \,.$$

Opmerking: Dit kan natuurlijk ook direct, daar $AX = Y$ betekent $Ax_{*j} = y_{*j}$ voor alle j!

## § 1. Enkele punten uit de matrixrekening

Algemeen

1.1. Als $A_{pq}B^-_{qp} = I_p$ en $p \neq q$, dan is $BA \neq I_q$.

1.2. Ga na dat $(A,B) := sp(A'B)$ inderdaad een inprodukt definieert.

1.3. Zij $\lambda$ eigenwaarde van een niet defekte $A_{nn}$. Bewijs:
de algebraïsche multipliciteit van $\lambda$ = meetkundige multipliciteit van $\lambda$;
zie ook (1.20).

1.4. Geef een zo eenvoudig mogelijk voorbeeld waarbij $r(AB) \neq r(BA)$.

1.5. Bewijs: $r(A+B) \leq r(A) + r(B)$.

1.6. Zij $A = \{a_{ij}\}$ met $\forall_i\ a_{ii} = a$ en $\forall_{i \neq j}\ a_{ij} = b$. Bereken $|A|$.

1.7. Bewijs:

$$\begin{vmatrix} A_{pp} & B \\ C & I_q \end{vmatrix} = |A - BC|.$$

1.8. In Rao: "Linear statistical inference and its applications" pag. 38 staat:
"The proof (of 1.83) consists in showing that A and B have the same eigenvectors". Dit is onjuist. Geef tegenvoorbeeld.

1.9. Het bewijs van (1.83) verloopt moeizaam. Geef een eenvoudig bewijs voor het geval dat alle $\lambda_i$ van B <u>verschillend</u> zijn.

1.10. $A \in M_{n\times n}$. Bewijs: $A'A = AA' \Rightarrow N(A) = N(A')$

$A,B \in M_{n\times n}$, symmetrisch. Bewijs: $A^2 = B^2 \Rightarrow N(A) = N(B)$.

1.11. A,B symmetrisch. Bewijs: $sp[(AB)^2] \leq sp(A^2B^2)$. (Hint, beschouw $(AB - BA)^2$).

Niet negatief definiete matrices

1.20. Construeer een (niet diagonale) positief definiete en semi positief definiete matrix van de orde 3.

1.21. Geef een voorbeeld waaruit blijkt dat (1.68) niet geldt als A s.p.d. is.

1.22. Is dit juist: $A,B,C > 0$: $[A > B \wedge B > C] \Rightarrow A > C$ ?

1.23. In (1.78) staat: $A > 0 \Rightarrow |A| \leq \Pi a_{ii}$. Wanneer geldt het gelijkteken ?

1.24. Bewijs: $\|A\| < 1 \Rightarrow I \pm A$ regulier (zie ook (1.77)).

1.25. Bewijs: $A, B > 0$ èn $AB$ symmetrisch $\Rightarrow AB > 0$

$\alpha,\beta > 0;\ A,B > 0 \quad \Rightarrow \alpha A + \beta B > 0$.

1.26. Zij $A, B > 0$ en $A \geq B$. Bewijs:

a) $\forall_i\ a_{ii} \geq b_{ii}$

b) $|A| \geq |B|$

c) $\operatorname{sp} A \geq \operatorname{sp} B$

d) $B^{-1} \geq A^{-1}$

e) $\|A\| \geq \|B\|$

f) $\forall C : C'AC \geq C'BC$

g) $\Sigma_{ij}\ a_{ij} \geq \Sigma_{ij}\ b_{ij}$

1.27. Geldt altijd $\|A_a\| = \|A\|$ ? Zo nee, geef tegenvoorbeeld.

1.28. Bewijs: $A > 0$; $BAB' = 0 \Leftrightarrow B = 0$.

1.29. Bewijs: $A > 0$; $r(BAB') = r(B)$.

1.30. Bewijs: $A > 0$; $\exists$ reguliere $B$ zodat $BAB' = I$.

1.31. Bewijs: $A > 0$; $|A| \leq |A_{11}||A_{22}|$. Wanneer geldt het gelijkteken ?

1.32. Bekend is: $f(X) = x^{-1}$ is voor $x > 0$ een konvexe funktie van x.

Bewijs: $f(X) = X^{-1}$ is voor $X > 0$ een konvexe funktie van X.

Dus te bewijzen: $\forall\ X,Y > 0$ en $0 \leq \alpha \leq 1$ geldt:

$$\alpha X^{-1} + (1 - \alpha)Y^{-1} \geq [\alpha X + (1 - \alpha)Y]^{-1}$$

1.33. Bewijs: $A,B \geq 0$; $A + B = 0 \Leftrightarrow A = B = 0$.

1.34. Bewijs: $A,B \geq 0$; $\operatorname{sp}(AB) = 0 \Leftrightarrow AB = 0$.

1.35. Bewijs: $A,B \geq 0,\ 0 \leq \alpha \leq 1$: $|\alpha A + (1 - \alpha)B| \geq |A|^{\alpha}|B|^{1-\alpha}$.

Hieruit volgt dat $\log|X|$ een concave funktie is van X.

1.36. Bewijs: $A,B > 0$ van orde n; $\operatorname{sp}(AB) \geq n|A|^{1/n}|B|^{1/n}$.

1.37. Zij $A = \frac{1}{3}\begin{pmatrix} 7 & -4 & 2 \\ -4 & 7 & -2 \\ 2 & -2 & 4 \end{pmatrix} > 0$. Bepaal een B en een $D \neq B$ zodat $A = B^2 = D^2$.

1.38. Zij $|A - \lambda I| = \sum_0^n (-1)^i p_i(A)\lambda^i$. Bewijs: $A > 0 \Leftrightarrow \forall_i p_i(A) > 0$.

1.39. Bewijs: $AXX' = 0 \Leftrightarrow AX = 0$.

1.40. Bewijs: $A = \{a_{ij}\} > 0 \Rightarrow A^{\star} := \{|a_{ij}|\} > 0$.

1.41. Zij $A$ symmetrisch; $B > 0$. Bewijs $\exists t$ zodat $B - tA > 0$ zie (4.10).

Idempotente matrices

1.50. Bewijs: $P := X(X'X)^{-1}X'$ is de orthogonale projector op $<X>$. (zie (1.91)).

1.51. Zij P een orthogonale projector. Bewijs: $\|Px\| \leq \|x\|$ en $\cos(Px,x) \geq 0$.

1.52. Bewijs: A'A orthogonale projector $\Leftrightarrow$ AA' orthogonale projector.

1.53. P orthogonale projector Bewijs: $\forall i \quad 0 \leq p_{ii} \leq 1$ en $\forall i \neq j \; |p_{ij}| \leq \frac{1}{2}$.

1.54. Zij $A_{nn}$ symmetrisch en $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ de geordende eigenwaarden van A.
Bewijs

a) $\max\limits_{x \neq 0} \frac{x'Ax}{x'x} = \lambda_1$ en $\min\limits_{x \neq 0} \frac{x'Ax}{x'x} = \lambda_n$

b) $\forall_i : \lambda_n \leq a_{ii} \leq \lambda_1$

c) $\lambda_n \leq \sum\limits_{ij} a_{ij}/n \leq \lambda_1$

d) Zij $B > 0$; $\max\limits_{x \neq 0} \frac{x'Ax}{x'Bx} = \lambda_1$ (= grootste eigenwaarde van $B^{-1}A$)

1.55. Zij A niet defekt. Als voor een $k \in \mathbb{N}$ geldt $A^k = A^{k+1}$, dan is A idempotent.
Bewijs dit en geef tegenvoorbeeld voor defekte A.

1.56. P,Q orthogonale projectoren. Bewijs: $sp(PQ) \leq \min(sp\,P, sp\,Q)$.
Als $sp(PQ) = sp\,P \Rightarrow PQ = P$. Ook geldt: $sp(PQ) = 0 \Leftrightarrow PQ = 0$.
Opmerking: $(P,Q) = 0$ is dus equivalent met $PQ = 0$.

1.57. $A_i$, $i = 1,\ldots,k$ zijn symmetrische $n \times n$-matrices. $A := \sum\limits_1^k A_i$.
Bewijs: elke twee van de drie voorwaarden a, b en c impliceren de derde:
a) A orthogonale projector.
b) elke $A_i$ orthogonale projector.
c) de $A_i$ zijn disjunct dat wil zeggen $\forall_{i \neq j} \; A_iA_j = 0$.

1.58. Bewijs: P orthogonale projector: $(PA)^2 \leq A^2$.

## § 2. Enkele begrippen uit de mathematische statistiek

2.1. Geef een eenvoudig voorbeeld van een tweedimensionale verdeling $f(x,y)$ waarbij $\rho(\underline{x},\underline{y}) = 0$, terwijl $\underline{x}$ en $\underline{y}$ toch stochastisch afhankelijk zijn.

2.2. Bewijs (2.34.). Bedenk dat $\rho$ zelf dus op het teken na invariant is.

2.3. In (2.36) is $\mathcal{E}(\underline{x}'A\underline{x})$ afgeleid.
Bewijs: $\mathcal{E}(\underline{x}'A\underline{y}) = sp(A\Sigma_{21}) + \mu_x'A\mu_y$.

2.4. Wat betekent het als a) var $\underline{x} = \Sigma^2$ singulier is?
b) $\Sigma^2$ diagonaal is;

2.5. var $\underline{x} = \sigma^2 I$; $\underline{y} = a'\underline{x}$ en $\underline{z} = b'\underline{x}$.
Bewijs: $\underline{y}$ en $\underline{z}$ ongecorreleerd $\Leftrightarrow a'b = 0$.

2.6. Stel $\underline{x}$ is 1 dimensionaal symmetrisch verdeeld rond 0. Bewijs: $\text{cov}(\underline{x},\underline{x}^2) = 0$.

## § 3. De multinormale verdeling.

3.1. Zij $\Sigma^2 > 0$. Bewijs: a) $\Sigma_{11}, \Sigma_{22}, \Sigma_{11.2}, \Sigma_{22.1} > 0$.

$$\text{b) } \Sigma_{1.2}, \Sigma_{2.1}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{2.1} \end{pmatrix} \geq 0.$$

Stel nu $\Sigma^2 \geq 0$, dan c) $\langle\Sigma_{12}\rangle \subset \langle\Sigma_{11}\rangle$

3.2. Zij $\Sigma^2 > 0$; dan is $\Sigma_{11} \geq \Sigma_{11.2}$ en $\Sigma_{11} > \Sigma_{1.2}$. Bewijs.

3.3. Bewijs: $|\Sigma_{11.2}| = |\Sigma_{11}| \Leftrightarrow \Sigma_{11.2} = \Sigma_{11} \Leftrightarrow \Sigma_{12} = 0$.

3.4. Bewijs: $|\Sigma^2| \leq |\Sigma_{11}||\Sigma_{22}$ . Het gelijkteken geldt d.e.s.d. als $\Sigma_{12} = 0$.

3.5. Uit vraagstuk 3.4 volgt eenvoudig (1.78). Ga na.

3.6. Op pag. 3.2 staat: "doorsnijding met vlak $f(x)$ = constant geeft ellipsen". Verklaar dit. Wanneer zijn het cirkels?

3.7. Bewijs theorema 3.9 met behulp van de kansdichtheid.

3.8. Zij $R$ een correlatiematrix met $\forall i \neq j$ $\rho_{ij} = \rho \neq 0$ $(i,j = 1,\dots,n)$. Bewijs $-1/(n-1) \leq \rho \leq 1$. Wanneer is $|R| = 0$?

3.9. Bewijs eveneens: $0 \leq |R| \leq 1$. Wanneer is $|R| = 1$ ?

3.10. In Graybill: "Introduction to matrices with applications in statistics", pag. 220 staat: "Show that the largest root of a correlationmatrix is less or equal unity".
Bewijs dat dit onjuist is.

3.11. Zij $\underline{x}$ multinormaal verdeeld met kansdichtheid $f(x)$.
Bewijs: $\underline{x}$ is die x-waarde waarvoor $\frac{df(x)}{dx} = 0$.

3.12. In welke van de volgende gevallen heeft elke lineaire kombinatie $\alpha'\underline{x}$ $(\alpha \neq 0 \in \mathbb{R}^p)$ van de komponenten van $\underline{x} = (\underline{x}_1,\dots,\underline{x}_p)'$ een normale verdeling ?

a) $\underline{x}_i \sim N(\mu_i,\sigma_i^2)$ $i = 1,\dots,p$

b) $\underline{x}_i \sim N(\mu_i,\sigma_i^2)$ $i = 1,\dots,p$ en de $\underline{x}_i$ zijn onderling ongecorreleerd.

c) $\underline{x}_i \sim N(\mu_i,\sigma_i^2)$ $i = 1,\dots,p$ en de $\underline{x}_i$ zijn onderling stochastisch onafhankelijk.

d) $\underline{x} \sim N_p(\mu,\Sigma^2)$

3.13. Zij A symmetrisch en B > 0. Bewijs:

$$I = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}(x'Ax + a'x + \alpha)\exp[-(x'Bx + b'x + \beta)]dx_1\ldots dx_n =$$

$$= \tfrac{1}{2}\pi^{n/2}|B|^{-\frac{1}{2}}\{sp(AB^{-1}) - a'B^{-1}b + \tfrac{1}{2}b'B^{-1}AB^{-1}b + 2\alpha\}\exp(\tfrac{1}{4}b'B^{-1}b - \beta).$$

Is AB > 0 dan geldt:

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}e^{-x'ABx}dx_1\ldots dx_n = \pi^{n/2}|AB|^{-\frac{1}{2}}.$$

3.14. Zij $\underline{x} \sim N_n(\mu,\sigma^2 I)$ met $\mu = (\mu,\ldots,\mu)'$

Bewijs:

$$\frac{\sqrt{n(n-1)}(\bar{\underline{x}} - \mu)}{\sqrt{\sum(\underline{x}_i - \bar{\underline{x}})^2}} \simeq \underline{t}_{n-1}.$$

3.15. Zij $\underline{x} \sim N_2(\mu,\Sigma^2)$ met $\mu_1 = \mu_2$ en $\sigma_1^2 = \sigma_2^2$.

Bewijs: $\underline{x}_1 + \underline{x}_2$ en $\underline{x}_1 - \underline{x}_2$ zijn stochastisch onafhankelijk.

3.16. Zij $\underline{x} \sim N(\mu,\sigma^2)$ en $\underline{x}_1,\ldots,\underline{x}_n$ een onafhankelijke steekproef van $\underline{x}$.
Definieer $\underline{y} = (\bar{\underline{x}},\underline{x}_i - \bar{\underline{x}})$. Bewijs dat $\underline{y} \sim N_2$ met var $\underline{y} = \frac{\sigma^2}{n}\begin{pmatrix}1 & 0\\ 0 & n-1\end{pmatrix}$.
(met andere woorden $\bar{\underline{x}}$ en $\underline{x}_i - \bar{\underline{x}}$ zijn stochastisch onafhankelijk en dus ook $\bar{\underline{x}}$ en $\underline{s}^2 = \sum(x_i - \bar{x})^2/(n-1)$)

§ 4. De verdeling van kwadratische vormen

4.1. $\underline{x} \sim N_p(\mu, I)$; $A_i$ $i = 1,\ldots,k$ zijn symmetrische $p \times p$ matrices met

$$\underline{x}'\underline{x} = \sum_1^k \underline{x}'A_i\underline{x}$$

Bewijs: elk van de voorwaarden

1) elke $A_i$ is een orthogonale projector

2) de $A_i$ zijn disjunct, is noodzakelijk en voldoende opdat:

a) $\underline{x}'A_i\underline{x} \simeq \underline{\chi}^2_{r_i}(\mu'A\mu)$ $i = 1,\ldots,k$ en $r_i = r(A_i)$

b) $\underline{x}'A_i\underline{x}$ onderling onafhankelijk zijn.

4.2. $\underline{y}_i \sim N(\mu_i, \sigma_i^2)$ $i = 1,\ldots,n$; $\underline{y}_1,\ldots,\underline{y}_n$ zijn stochastisch onafhankelijk.

Definieer $\underline{y}^* := \sum \sigma_i^{-2}\underline{y}_i / \sum \sigma_i^{-2}$

Bewijs: $\sum \sigma_i^{-2}(\underline{y}_i - \underline{y}^*)^2 \simeq \underline{\chi}^2_{n-1}(\delta)$ met $\delta = 0$ als $\mu_1 = \ldots = \mu_n$.

4.3. Zij $\underline{x} \sim N_p(\mu_1, \Sigma_{11})$, $\underline{y} \sim N_q(\mu_2, \Sigma_{22})$ en $(\underline{x},\underline{y}) \sim N_{p+q}(\mu, \Sigma^2)$.

$A \in M_{pq}$. Probeer de bilineaire vorm $\underline{x}'A\underline{y}$ te schrijven als een kwadratische vorm $\underline{z}'B\underline{z}$. Bepaal nu $\mathcal{E}(\underline{x}'A\underline{y})$ en $\mathrm{var}(\underline{x}'A\underline{y})$.

4.4. Als $\underline{x} \sim N_p(\mu, \Sigma^2)$, bewijs dan:

$$\mathrm{cov}(\underline{x}'A\underline{x}, \underline{x}'B\underline{x}) = 2\ \mathrm{sp}(A\Sigma^2B\Sigma^2) + 4\mu'A\Sigma^2B\mu .$$

Voor $A = B$ onstaat formule (4.7).

4.5. Zij $\underline{x} \sim N_2(0, \Sigma^2)$. Bewijs: $\rho(\underline{x}_1^2, \underline{x}_2^2) = \rho^2(\underline{x}_1, \underline{x}_2)$.

4.6. $[VAVAV = VAV$ en $\mu'AV = \mu'AVAV] \Leftrightarrow [VAVAV = VAV$ en $VA\mu \in \langle VAV \rangle]$

Bewijs dit. (A,V zijn symmetrisch).

4.7. Zij $V \geq 0$, A symmetrisch. Bewijs: $VAVAV = VAV \Leftrightarrow (VA)^3 = (VA)^2$.

§ 5. Regressie, het model van volle rang

5.1. Zij $\underline{y} = X\beta + \underline{e}$; X volle kolomrang, $\mathcal{E}\underline{e} = 0$, var $\underline{e} = \sigma^2 I$.

De kleinste kwadratenschatter (afgekort k.k.s) voor $\beta$ is $\underline{b} = (X'X)^{-1}X'\underline{y}$ (zie pag. 5.5).

Bewijs: elke andere zuivere lineaire schatter $\tilde{\underline{b}}$ is "slechter", d.w.z. VAR $\tilde{\underline{b}} \geq$ VAR $\underline{b}$.

5.2. Analoog aan 5.1, alleen nu VAR $\underline{e} = \Sigma^2 > 0$ ondersteld.

De gewogen k.k.s. voor $\beta$ is $\underline{b} = (X'\Sigma^{-2}X)^{-1}X'\Sigma^{-2}\underline{y}$ (zie pag. 5.6).

Bewijs: elke andere zuivere lineaire schatter is "slechter".

5.3. Zij $\underline{y} = X\beta + \underline{e}$; X volle kolomrang; $\mathcal{E}\underline{e} = 0$, VAR $\underline{e} = \sigma^2 I$.

$\hat{\underline{y}} = X\underline{b}$. Bewijs VAR $\hat{\underline{y}} \leq$ VAR $\underline{y}$.

5.4. Analoog aan som 5.3, alleen nu var $\underline{e} = \Sigma^2 > 0$ ondersteld.

$\hat{\underline{y}} = X\underline{b}$ met $\underline{b}$ gewogen k.k.s. Bewijs: VAR $\hat{\underline{y}} \leq$ VAR $\underline{y}$.

5.5. Weer hetzelfde model met var $\underline{e} = \Sigma^2 > 0$. Bewijs direkt of met behulp van som 5.2 dat de gewogen k.k.s. voor $\beta$ "beter" is (kleinere variantie heeft) dan de ongewogen k.k.s.

5.6. Beschouw het model $\underline{y} = X\beta + \underline{e}$ met $\underline{e} \sim N_n(0,\sigma^2 I)$.

Vindt de maximum likelihoodschatter $\hat{\underline{\sigma}}^2$ voor $\sigma^2$. Deze is niet zuiver, doch heeft een kleinere variantie dan $\underline{s}^2 = SSE/(n - r)$.

5.7. Bewijs: $X_m'y_m = X_m'y$, zie (5.16).

5.8. In het model $\underline{y} = X\beta + \underline{e}$ met $\beta = (\beta_0,\beta_1,\beta_2)'$ zijn de normaalvergelijkingen als volgt:

$$3b_0 + b_1 - 2b_b = 1$$

$$b_0 + 2b_1 + b_2 = 7 \qquad \text{Bereken } b := (b_0,b_1,b_2)' \text{ en } \|\hat{y}\|^2.$$

$$-2b_0 + b_1 + 4b_2 = 9$$

5.9. Beschouw het eenvoudige lineaire model van pag. 5.2.

$\underline{y}_i = \beta_0 + \beta_1 x_i + \underline{e}_i \qquad i = 1,\ldots,n;\ \underline{e} \sim N_n(0,\sigma^2 I)$.

a) Bepaal via de algemene formules de schatter $\underline{b} = (\underline{b}_0,\underline{b}_1)'$ en tevens var $\underline{b}$.

b) Kies de $x_i$ zodanig dat var $\underline{b}_0$ en var $\underline{b}_1$ minimaal zijn. Bewijs dat dan $\text{cov}(\underline{b}_0,\underline{b}_1) = 0$.

c) Bewijs dat $s^2 = \frac{1}{n-2}[\sum(y - \bar{y})^2 - \frac{\{\sum(x-\bar{x})(y-\bar{y})\}^2}{\sum(x-\bar{x})^2}]$.

5.10. In het model $\underline{y} = X\beta + \underline{e}$, $\underline{e} \sim N_n(0,\sigma^2 I)$ definiëren we $\hat{y} = P_x y$, $\bar{y} = Py$, $\hat{y}_m = P_{x_m} y$, de orthogonale projecties van y respektievelijk op $\langle X \rangle$, $\langle u \rangle$ en $\langle X_m \rangle$. Bewijs:

a) $\langle X \rangle = \langle X_m\ u \rangle$; $u \perp \langle X_m \rangle$

b) $P_{x_m} = P_x - P$; $P_x P_{x_m} = P_{x_m} P_x = P_{x_m}$; $P_x P = PP_x = P$; $P_{x_m} P = PP_{x_m} = 0$.

## § 6. Verdelingen in het regressiemodel

6.1. Bewijs dat de som van de niet-centraliteitsparameters van de verdelingen van SSM, $SSR_m$ en SSE gelijk is aan die van de verdeling van SST.

6.2. Analoog aan $\hat{y} = P_x y = Xb$ geldt: $\hat{y}_m := P_{x_m} y_m = X_m b_m$ (zie 6.6 ci).

6.3. Gegeven is het lineaire model: $\underline{Y} = \beta_0 + \beta_1 x + \beta_2 x^2 + \underline{e}$, $\underline{e} \sim N(0,\sigma^2)$

en de waarnemingen

| x | -2 | -1 | 0 | 1 | 2 |
|---|---|---|---|---|---|
| y | -7 | -2 | -1 | 1 | 9 . |

a) Bereken de kleinste kwadratenschattingen $b_0$, $b_1$ en $b_2$.

b) Toets m.b.v. een betrouwbaarheidsinterval de hypothese $H_0 : \beta_2 = 0$, $(\alpha = 0.05)$

c) Geef een betrouwbaarheidsinterval voor $\beta_1$ (95% interval).

6.4. Gegeven is het lineaire model: $\underline{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \underline{e}$.
Er zijn in de volgende 9 punten totaal 13 waarnemingen verricht als volgt:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $x_1$: | -1 | 1 | -1 | 1 | $-\sqrt{2}$ | $\sqrt{2}$ | 0 | 0 | 0 |
| $x_2$: | -1 | -1 | 1 | 1 | 0 | 0 | $-\sqrt{2}$ | $\sqrt{2}$ | 0 |
| y : | 9,16 | 13,37 | 7,82 | 10,35 | 9,32 | 13,62 | 10,51 | 8,84 | 10,64; 10,11; 9,98; 9,14; 10,01 |

Dit zijn dus 8 waarnemingen in punten gelegen op een cirkel, en 5 waarnemingen in de oorsprong.

a) Schat de parameters en $\sigma^2$ (ondersteld $\underline{e} \sim N(0,\sigma^2)$).

b) Toets de hypothese dat het model juist is door SSE te splitsen in SSLF en SSPE (zie pag. 6.4).

6.5. In (6.8) staat

$$\frac{\|\hat{\underline{y}} - n\|^2/r}{\|\hat{\underline{y}} - \underline{y}\|^2/(n-r)} = \underline{F}^r_{n-r} .$$

Dit is bewezen met behulp van (6.4) $\underline{s}^2$ onafhankelijk $\underline{b}$.

Bewijs dit direkt, dus dat de teller een centrale $\underline{\chi}^2_r$-, en de noemer een $\underline{\chi}^2_{n-r}$- verdeling heeft en beide onafhankelijk!

## § 7. De algemene lineaire hypothese

7.1. Bewijs: $\hat{y}_{x_H} \perp \hat{y} - \hat{y}_{x_H}$ (zie pag. 7.5).

7.2. Beschouw model en waarnemingen uit vraagstuk 6.4.
Toets de hypothese dat de 2e graadstermen weggelaten kunnen worden.

7.3. Toets de hypothese $H_0 : \beta_2 = 0$ uit vraagstuk 6.3 b), maar nu met behulp van $K\beta = \gamma$. Bepaal de schatting voor $\beta$ onder $H_0$, $b_{x_H}$

7.4. Beschouw het model $\underline{y} = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \underline{e}$, $\quad \underline{e} \sim N(0,\sigma^2)$

Gegeven zijn de waarnemingen

| $x_1$ | $x_2$ | $x_3$ | $y$ |
|---|---|---|---|
| 2 | 1 | 4 | 8 |
| -1 | 2 | 1 | 10 |
| 1 | -3 | 4 | 9 |
| 2 | 1 | 2 | 6 |
| 1 | 4 | 6 | 12 |

a) Bepaal de schatting b.

b) Toets de hypothese $H_0$: $\beta_1 = \beta_2 + 4$.

c) Bepaal $b_{x_H}$.

7.5. Bewijs dat $P_H := XX^{\underline{-2}}K'(KX^{\underline{-2}}K')^{-1}KX^{\underline{-2}}X'$ een orthogonale projector is op $H := \langle XX^{\underline{-2}}K' \rangle$ (zie pag. 7.4 onderaan).

7.6. Op pag. 7.5 staat: $\hat{y}_{x_H} = P_{x_H} y = P_x y - P_H y = \hat{y} - P_H y$.
Echter op pag. 7.4 naast tabel staat gedefinieerd: $\hat{y}_{x_H} = \hat{y} - \hat{y}_H$.
Bewijs dus dat $P_H y = \hat{y}_H$.

7.7. $P_x$ en $P_H$ zijn de orthogonale projectoren op $\langle X \rangle$ resp. $\langle H \rangle := \langle XX^{\underline{-2}}K' \rangle$.
Bewijs: $P_{x_H} := P_x - P_H$ is de orthogonale projector op $X_H := \langle X \rangle \cap \langle H \rangle^{\perp}$.

8.1. De volgende waarnemingen zijn uitgevoerd bij behandelingen volgens een zogenaamd "Latijns vierkant".

waarnemingen ($i$ = rij, $j$ = kolom):

| $i$ \ $j$ | | | |
|---|---|---|---|
| | 8 | 3 | 12 |
| | 4 | 11 | 7 |
| | 9 | 7 | 10 |

behandelingen.

| | | |
|---|---|---|
| 1 | 2 | 3 |
| 2 | 3 | 1 |
| 3 | 1 | 2 |

Beschouw het model $\underline{y}_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \underline{e}_{ijk}$; $i,j,k = 1,2,3$.

Toets de hypothese a) $\alpha_1 = \alpha_2 = \alpha_3$ b) $\beta_1 = \beta_2 = \beta_3$ c) $\gamma_1 = \gamma_2 = \gamma_3$

met andere woorden of er verschil is tussen resp. rijen, kolommen, behandelingen.

8.2. Gegeven de volgende waarnemingen, volgens een "genest model"

$\underline{y}_{ijk} = \mu + \alpha_i + \beta_{ij} + \underline{e}_{ijk}$.

a) Geef een schatting $b^0$ voor $\beta$

b) Geef een variantie-analyse

c) Toets de $H_0$: er is geen verschil in secties binnen de cursus.

| i cursus: | Engels | | Aardrijkskunde | | |
|---|---|---|---|---|---|
| j sectie | 1 | 2 | 1 | 2 | 3 |
| | 5 | 8 | 8 | 6 | 3 |
| k ↓ | | 10 | 10 | 2 | 7 |
| | | 9 | | 1 | |
| | | | | 3 | |

9.1. Gegeven is het volgende lineaire model:

$$\underline{y}_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \underline{e}_{ijk}; \quad i = 1,2;\ j = 1,2;\ k = 1,2.$$

Ga na welke van de volgende funkties schatbaar zijn:

a) $\mu + \alpha_1$ b) $\mu + \alpha_1 + \beta_1$ c) $\mu + \alpha_1 + \beta_1 + \gamma_{11}$ d) $\alpha_1 - \alpha_2$ e) $\alpha_1 - \alpha_2 + \gamma_{11} - \gamma_{2}$

9.2. a) Geef van vraagstuk 8.2 alle schatbare funkties weer in 1 vorm.

b) Ga na dat $\alpha_1 - \alpha_2$ niet schatbaar is.

$\alpha_1 - \alpha_2$ is wèl schatbaar onder de aanname (zoals in een variantie-analyse)

$\sum_j \beta_{1j} = \sum_j \beta_{2j} = 0$ (d.w.z. in het beperkte model).

10.1. Ga na of de hypothesen in vraagstuk 8.1 inderdaad toetsbaar zijn.

10.2. Is in vraagstuk 8.2 c) de hypothese toetsbaar?

12.1. Ga na dat in (12.5) voor V = I de bekende relaties ontstaan namelijk

$$Kb^0 = KGX'y \ (10.5), \ \mathrm{VAR}\ K\underline{b}^0 = \sigma^2 KGK' \ (9.7) \text{ en } s^2 = (\underline{y}^2 - \underline{\hat{y}}^2)/\nu .$$