# Matrix Algebra med Excel
# Forelæsningsnoter til FR86

Jesper Lund

mail@JesperLund.com
http://www.JesperLund.com

28. august 2002

# 1 Indledning

Matrix algebra er et uundværligt redskab til økonometri, herunder fag som empirisk finansiering. Formålet med denne note er at beskrive de basale regneregler for matrix algebra, svarende nogenlunde til hvad der forudsættes ved læsning af lærebogen for faget FR86 [Campbell, Lo and MacKinlay (1997)]. Noten beskriver desuden hvordan regnearket Microsoft Excel kan anvendes til (visse) matrix algebra operationer.

# 2 Definitioner

## 2.1 Matrix

En matrix er en rektangulær tabel af tal (rektangulær vil sige at hver række og hver søjle indeholder lige mange tal). Notationen en "$n \times m$ matrix" (læses: $n$ kryds $m$) betyder en matrix med $n$ rækker og $m$ søjler (kolonner). Elementet i den $i$'te række og $j$'te søjle benævnes $a_{ij}$, og hele matricen skrives som

$$\mathbf{A}_{n\times m} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{22} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \tag{1}$$

eller mere kortfattet $\mathbf{A} = [a_{ij}]$. Notationen $\cdots$ betyder at vi har udeladt et vilkårligt antal rækker eller søjler ved opskrivningen. Bemærk desuden at man normalt skriver matricer med stort $\mathbf{A}$, mens de enkelte elementer skrives med lille $a_{ij}$ efterfulgt af de relevante fodtegn. Ligeledes skrives matricer (og vektorer) ofte med boldface skrifttype, sådan som det er gjort her.

Et mere konkret eksempel er en $2 \times 3$ matrix, der har følgende elementer

$$\mathbf{A}_{2\times 3} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \tag{2}$$

Ofte udelader man fodtegnet $n \times m$, der angiver dimensionerne på matricen, men i starten når man lærer matrix algebra, er det en god ide konsekvent at skrive dimensionerne på matricen, blandt andet fordi regneregler som matrix multiplikation kun gælder hvis dimensioner overholder visse betingelser (mere herom senere).

En $n \times m$ matrix siges at være kvadratisk hvis $n = m$. For en kvadratiske matrix $\mathbf{A}_{n\times n}$ kalder man elementerne $a_{ii}$ for diagonalelementerne.

## 2.2 Transponering

Transponering af matricen $\mathbf{A}$ benævnes ved $\mathbf{A}'$ (læses: $\mathbf{A}$ mærke), og det vil sige at man bytter om på rækker og søjler. Hvis man f.eks. transponerer matricen $\mathbf{A}_{2\times 3}$ fås

$$\mathbf{A}'_{3\times 2} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix} \tag{3}$$

og for den generelle $n \times m$ matrix giver transponering

$$\mathbf{A}'_{m\times n} = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{bmatrix}. \tag{4}$$

## 2.3 Lighed mellem matricer og symmetri

To matricer $\mathbf{A}$ og $\mathbf{B}$ siges at være lig med hinanden, hvis de har samme dimensioner, og hvis $a_{ij} = b_{ij}$ for samtlige elementer i de to matricer.

En matrix $\mathbf{A}$ kaldes symmetrisk, hvis $\mathbf{A} = \mathbf{A}'$. Dertil kræves at matricen er kvadratisk, og at $a_{ij} = a_{ji}$ for alle $i \neq j$.

## 2.4 Vektor

En vektor er et specialtilfælde af en matrix, hvor der kun er en søjle.

$$\mathbf{v}_{n\times 1} = \begin{bmatrix} v_1 \\ v_2 \\ \cdots \\ v_n \end{bmatrix} \tag{5}$$

Strengt taget burde man kalde ovennævnte en søjlevektor. Der findes nemlig også rækkevektorer, som er matricer med en række og $n$ søjler. Når man blot skriver "vektor" mener man normalt en søjlevektor. Transponering gælder også for vektorer, dvs. hvis $\mathbf{v}$ er en (søjle)vektor, så er $\mathbf{v}'$ en rækkevektor.

Nogle gange er der hensigtsmæssigt at skrive en matrix vha. søjle- eller rækkevektorer. For eksempel kan den generelle $n \times m$ matrix skrives som

$$\mathbf{A}_{n\times m} = \begin{bmatrix} \mathbf{a}'_1 \\ \mathbf{a}'_2 \\ \cdots \\ \mathbf{a}'_n, \end{bmatrix} \tag{6}$$

hvor $\mathbf{a}_k$ er en (søjle)vektor med $m$ elementer (således at $\mathbf{a}'_k$ er en $1 \times m$ rækkevektor). Her angiver fodtegnet $k$ ikke dimensionen af vektoren, men placeringen (rækken) i matricen $\mathbf{A}$.

## 2.5 Nogle specielle matricer og vektorer

Identitetsmatricen $\mathbf{I}_{n\times n}$ er en symmetrisk matrix, der har 1 på diagonalen og 0 alle andre steder. For den generelle $\mathbf{I}_{n\times n}$ identitetsmatrix har vi

$$\mathbf{I}_{n\times n} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}. \tag{7}$$

En diagonal matrix er en kvadratisk matrix, hvor alle elementer uden for diagonalen er 0. Det vil sige at hvis $\mathbf{D}$ er en diagonal matrix, er $d_{ij} = 0$ for alle $i \neq j$. Identitetsmatricen er således et specialtilfælde af en diagonal matrix, hvor alle diagonalelementerne er 1.

Ettalsvektoren er en vektor hvor alle elementer er 1. Ofte benævnes denne vektor ved symbolet $\boldsymbol{\iota}$, og vi har således $\boldsymbol{\iota}_{n\times 1} = (1 \ 1 \ \cdots \ 1)'$. Bemærk at $\boldsymbol{\iota}_{n\times 1}$ er en søjlevektor, men at jeg af pladshensyn har skrevet den som en rækkevektor og derefter transponeret den.

# 3 Regneregler for matricer

De følgende regneregler for matricer gælder også for vektorer, idet man jo skal huske at en $n$-dimensional vektor er en $n \times 1$ matrix.

## 3.1 Addition og subtraktion af matricer

Addition mellem to matricer defineres som elementvis addition, og operationen er derfor kun defineret hvis de to matricer har samme dimensioner (sagt formelt: er konforme for addition). Hvis $\mathbf{A}$ og $\mathbf{B}$ begge er $n \times m$ matricer, er

$$\mathbf{C}_{n\times m} = \mathbf{A}_{n\times m} + \mathbf{B}_{n\times m} \tag{8}$$

en $n \times m$ matrix med elementerne $c_{ij} = a_{ij} + b_{ij}$.

Subtraktion defineres på samme måde som addition (elementvis subtraktion), dvs. elementerne i $\mathbf{C} = \mathbf{A} - \mathbf{B}$ er givet ved $c_{ij} = a_{ij} - b_{ij}$.

Ved addition og subtraktion er rækkefølgen ligegyldig. Det betyder for eksempel at (bemærk brugen af paranteser)

$$\mathbf{D} = \mathbf{A} - (\mathbf{B} + \mathbf{C}) = \mathbf{A} - \mathbf{C} - \mathbf{B}. \tag{9}$$

Dette følger umiddelbart af definitionen på matrix addition som elementvis addition, hvorved vi har at $c_{ij} = a_{ij} - (b_{ij} + c_{ij})$.

## 3.2 Det indre produkt for vektorer

Det indre produkt mellem to $n$-dimensionale vektorer $\mathbf{a}$ og $\mathbf{b}$ benævnes $\mathbf{a}'\mathbf{b}$, og det er defineret som

$$\mathbf{a}'\mathbf{b} \ = \ \sum_{i=1}^{n} a_i b_i \ = \ a_1 b_1 + a_2 b_2 + \cdots + a_n b_n. \tag{10}$$

Vektorerne skal have samme dimension for at det indre produkt er defineret. Bemærk at denne regneregel kun er defineret for vektorer.

I visse sammenhænge taler man om "længden" af en vektor, benævnt ved notationen $\|\mathbf{v}\|$. Længden af vektoren $\mathbf{v}$ er defineret som $\|\mathbf{v}\| = \sqrt{\mathbf{v}'\mathbf{v}}$.

## 3.3 Matrix multiplikation

Hvis $\mathbf{A}_{n\times m}$ er en $n \times m$ matrix og $\mathbf{B}_{m\times K}$ er en $m \times K$ matrix, kan vi multiplicere $\mathbf{A}$ med $\mathbf{B}$. Matrix multiplikation defineres som det indre produkt at rækkerne i den første matrix med søjlerne i den anden. Det vil sige at elementerne i matrix multiplikationen

$$\mathbf{C}_{n\times K} = \mathbf{A}_{n\times m}\mathbf{B}_{m\times K} \tag{11}$$

givet ved

$$c_{ik} = \sum_{j=1}^{m} a_{ij} b_{jk}. \tag{12}$$

Bemærk at matrix multiplikationen $\mathbf{AB}$ er kun defineret hvis antallet af søjler i $\mathbf{A}$ svarer til antallet af rækker i $\mathbf{B}$ (hvorved matricerne er konforme for multiplikation). For at være helt præcis siger man at $\mathbf{A}$ *post*multipliceres med $\mathbf{B}$, eller at $\mathbf{B}$ *præ*multipliceres med $\mathbf{A}$.

Hvis man transponerer et matrix produkt $\mathbf{AB}$, gælder følgende regneregel:

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'. \tag{13}$$

For at indse dette, kan man betragte element $c'_{ji}$ i $\mathbf{C}'_{K\times n} = \mathbf{B}'_{K\times m}\mathbf{A}'_{m\times n}$, der kan skrives som

$$c'_{ji} = \sum_{k=1}^{K} b'_{jk} a'_{ki} = \sum_{k=1}^{K} a_{ik} b_{kj}, \tag{14}$$

idet $a'_{ki} = a_{ik}$. Den sidste sum i (14) er netop udtrykket (12) for element $c_{ij}$ i matrix produktet $\mathbf{C} = \mathbf{AB}$.

Hvis ellers dimensionerne passer, kan man multiplicere flere matricer med hinanden. For eksempel er følgende multiplikation defineret

$$\mathbf{D}_{n\times m} = \mathbf{A}_{n\times M}\mathbf{B}_{M\times K}\mathbf{C}_{K\times m}. \tag{15}$$

Det er underordnet om man først multiplicerer $\mathbf{A}$ med $\mathbf{B}$ eller $\mathbf{B}$ med $\mathbf{C}$. Derimod gælder det generelt at $\mathbf{AB} \neq \mathbf{BA}$, selvom begge matrix multiplikationer er definerede.

## 3.4 Skalar multiplikation

En skaler er et almindeligt tal (som et element i en matrix). Skalar multiplikationen $c\mathbf{A}$, hvor $c$ er en skalar, vil sige at alle elementer i matricen $\mathbf{A}$ multipliceres med $c$.

$$c\mathbf{A}_{n\times m} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1m} \\ ca_{22} & ca_{22} & \cdots & ca_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ ca_{n1} & ca_{n2} & \cdots & ca_{nm} \end{bmatrix} \tag{16}$$

Skalar multiplikation ændrer naturligvis ikke på dimensionerne af matricen.

### 3.5 Matrix determinant

Matrix determinanten er kun defineret for kvadratiske matricer $(n \times n)$. Man bruger notationen $|\mathbf{A}|$ eller $\det(\mathbf{A})$ for determinanten af matricen $\mathbf{A}$. For en $2 \times 2$ matrix $\mathbf{A}_{2\times 2}$ er determinanten givet ved udtrykket

$$|\mathbf{A}| = a_{11}a_{22} - a_{21}a_{12}. \tag{17}$$

For en generel $n \times n$ kan man skrive determinanten ved hjælp af den såkaldte co-faktor expansion. Ud fra en vilkårlig række $i$ kan man beregne determinanten som

$$|\mathbf{A}| = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}|\mathbf{A}_{ij}| = \sum_{j=1}^{n} a_{ij}C_{ij}, \tag{18}$$

hvor $\mathbf{A}_{ij}$ er den matrix der fremkommer ved at slette række $i$ og søjle $j$, og skalaren $C_{ij} = (-1)^{i+j}|\mathbf{A}_{ij}|$ kaldes co-faktoren. I ligning (18) skal vi altså beregne $n$ determinanter for matricer af dimensionen $(n-1) \times (n-1)$. Hvis man nu fortsætter med en co-faktor expansion for disse $n$ determinanter, vil man på et eller andet tidspunkt ende med en $2 \times 2$ matrix, hvis determinant kan beregnes ud fra formel (17). I praksis vil man naturligvis foretage determinant beregninger udover $3 \times 3$ ved hjælp af en computer, f.eks. regnearket Excel, mere herom i afsnit 4.

Hvis $\mathbf{A}$ og $\mathbf{B}$ begge er kvadratiske matricer, kan det vises at

$$|\mathbf{AB}| = |\mathbf{A}| \cdot |\mathbf{B}|. \tag{19}$$

Dette resultat betyder endvidere at

$$|\mathbf{AB}| = |\mathbf{BA}|, \tag{20}$$

selvom det generelt gælder at $\mathbf{AB} \neq \mathbf{BA}$.

For en diagonal matrix $\mathbf{D}_{n\times n}$ kan determinanten umiddelbart beregnes som

$$|\mathbf{D}| \;=\; \prod_{i=1}^{n} d_{ii} \;=\; d_{11}d_{22}\cdots d_{nn}. \tag{21}$$

Dette er let at se ved at bruge co-faktor expansionen ovenfor. En anden nyttig regneregel er at

$$|\mathbf{A}'| = |\mathbf{A}|, \tag{22}$$

dvs. at matricerne $\mathbf{A}$ og $\mathbf{A}'$ (begge $n \times n$) har samme determinant.

### 3.6 Matrix inversion

Ligesom determinanten for en matrix, gælder denne regneregel kun for kvadratiske matricer. Den inverse matrix af $\mathbf{A}$ er den matrix som opfylder betingelsen

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}, \tag{23}$$

hvor $\mathbf{I}$ er identitetsmatricen (der har 1 på diagonalen og 0 for alle andre elementer). Den inverse matrix er kun defineret for ikke-singulære matricer, og det kan vises at det svarer til at determinanten $|\mathbf{A}|$ skal være forskellig fra 0.

For en $2 \times 2$ matrix er den inverse defineret ved

$$\mathbf{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{21}a_{12}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \tag{24}$$

Bemærk at nævneren i skalaren er determinanten for $\mathbf{A}$, hvilket forklarer hvorfor den inverse kun er defineret hvis $|\mathbf{A}| \neq 0$. For den generelle $n \times n$ matrix kan det vises at

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{21} & \cdots & \mathbf{C}_{n1} \\ \mathbf{C}_{12} & \mathbf{C}_{22} & \cdots & \mathbf{C}_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{C}_{1n} & \mathbf{C}_{2n} & \cdots & \mathbf{C}_{nn} \end{bmatrix}, \tag{25}$$

hvor $C_{ij} = (-1)^{i+j}|\mathbf{A}_{ij}|$ er den $ij$'te co-faktor, altså $(-1)^{i+j}$ gange determinanten af den matrix, som fremkommer ved at slette række $i$ og søjle $j$ fra matricen $\mathbf{A}$. Ligesom for determinant beregninger er denne formel ikke videre praktisk for dimensioner udover $3 \times 3$. Computere løser heldigvis dette problem for os.

Følgende regneregler for $\mathbf{A}^{-1}$ er ofte nyttige

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A} \tag{26}$$

$$(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})' \tag{27}$$

$$|\mathbf{A}^{-1}| = \frac{1}{|\mathbf{A}|}. \tag{28}$$

Hvis $\mathbf{A}$ er symmetrisk, vil den inverse også være symmetrisk. Dette følger umiddelbart af ligning (27). Endelig gælder det at

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}, \tag{29}$$

forudsat at den inverse er defineret (kvadratisk dimension) og eksisterer for såvel $\mathbf{A}$ som $\mathbf{B}$ (svarende til at $|\mathbf{A}| \neq 0$ og $|\mathbf{B}| \neq 0$).

## 3.7 Trace af en matrix

Trace operatoren $\mathrm{Tr}(\mathbf{A})$ er kun defineret for kvadratiske matricer. Her gælder det at

$$\mathrm{Tr}(\mathbf{A}_{n\times n}) = \sum_{i=1}^{n} a_{ii}. \tag{30}$$

Trace af en matrix er altså summen af diagonalelementerne.

### 3.8 Kvadratiske former

Den kvadratiske form for en $n \times n$ symmetrisk matrix $\mathbf{C}$ er en skalar, som er defineret ved udtrykket [hvor $\mathbf{x}$ er en $n$-dimensional vektor]

$$\mathbf{x}'_{1\times n}\mathbf{C}_{n\times n}\mathbf{x}_{n\times 1} = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j c_{ij}. \tag{31}$$

Den kvadratiske form bruges blandt andet til at defineret egenskaben "positiv semi-definit" ved en matrix. En symmetrisk matrix $\mathbf{C}$ siges nemlig at være positiv semi-definit, hvis $\mathbf{x}'C\mathbf{x} \geq 0$ for alle værdier af vektoren $\mathbf{x}$. Hvis uligheden gælder som en streng ulighed, $\mathbf{x}'C\mathbf{x} > 0$, siges matricen $\mathbf{C}$ at være positiv definit. Det kan vises at kovariansmatricer altid er mindst positiv semi-definit, mere herom senere i afsnit 6.1.

Der findes forskellige metoder til at checke om en matrix er positiv definit eller ej, men de falder udenfor rammerne af denne forelæsningsnote.

### 3.9 Kronecker produkt[1]

Hvis vi har to matricer $\mathbf{A}_{n\times m}$ og $\mathbf{B}_{L\times K}$, kan vi definere det såkaldte Kronecker produkt $\mathbf{A} \otimes \mathbf{B}$ som en $Ln \times Km$ matrix med elementerne

$$(\mathbf{A} \otimes \mathbf{B})_{Ln\times Km} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1m}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2m}\mathbf{B} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1}\mathbf{B} & a_{n2}\mathbf{B} & \cdots & a_{nm}\mathbf{B} \end{bmatrix}. \tag{32}$$

Kronecker produktet er en $Ln \times Km$, der består af $nm$ submatricer, hver af dimensionen $L \times K$. Hver af disse submatricer fremkommer ved skalarmultiplikation mellem $a_{ij}$ og $\mathbf{B}$. En matrix bestående af submatricer kaldes en partitioneret matrix.

# 4 Excel funktioner til matrix algebra

Regnearket Excel har fire funktioner, som giver mulighed at lave basale matrix beregninger. Funktionsnavnene nedenfor refererer til den engelske udgave af Excel, og Microsoft har muligvis valgt andre navne i den danske udgave af Excel (i så fald må man konsultere `F1`-tasten, eller gætte sig frem).

- En matrix defineres som et range i Excel
- Funktionen `transpose` kan bruges til at beregne den transponerede matrix.
- Funktionen `mmult` multiplicerer to matricer med hinanden.
- Funktionen `mdeterm` beregner determinanten af en matrix.
- Funktionen `minverse` beregner den inverse matrix.

[1]Dette afsnit kan eventuelt overspringes uden tab af kontinuitet. Bemærk dog at Kronecker produktet bruges i Appendix A af Campbell et al. (1997)

Funktionerne `transpose`, `mmult` og `minverse` er array funktioner. For at se det fulde output, skal man markere et range af passende dimension (her regnereglerne for matrix algebra nyttige), og trykke `F2` og derefter `SHIFT-CTRL-ENTER` på samme tid.

# 5 Regneregler for vektor differentiation

## 5.1 Definition af begrebet vektor differentiation

Ved en "funktion af en vector" forstås en funktion af elementerne i vektoren:

$$F(\mathbf{x}) = F(x_1, x_2, \dots, x_n), \tag{33}$$

hvor $\mathbf{x}$ er en $n$-dimensional vektor. Ligning (33) kaldes også en funktion af $n$ variable.

Differentiation af $F(\mathbf{x})$ med hensyn til vektoren $\mathbf{x}$ er en $n$-dimensional vektor, hvor vi samler de $n$ partialle differentialekvotienter:

$$\frac{\partial F(\mathbf{x})}{\partial \mathbf{x}} = \begin{bmatrix} \dfrac{\partial F(x_1, x_2, \dots, x_n)}{\partial x_1} \\ \dfrac{\partial F(x_1, x_2, \dots, x_n)}{\partial x_2} \\ \cdots \\ \dfrac{\partial F(x_1, x_2, \dots, x_n)}{\partial x_n} \end{bmatrix} \tag{34}$$

Notationen $\frac{\partial F(\mathbf{x})}{\partial \mathbf{x}'}$ betyder at de $n$ partialle afledede stilles op som en $1 \times n$ række vektor, svarende til at $\frac{\partial F(\mathbf{x})}{\partial \mathbf{x}'} = \left(\frac{\partial F(\mathbf{x})}{\partial \mathbf{x}}\right)'$ (dette er dog mere notation end en egentlig regneregel).

Ligning (34) kaldes ofte gradienten af funktionen $F(\mathbf{x})$. Hessian matricen for funktionen $F(\mathbf{x})$ er defineret ved

$$\frac{\partial^2 F(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}'} = \begin{bmatrix} \dfrac{\partial^2 F(\mathbf{x})}{\partial x_1^2} & \dfrac{\partial^2 F(\mathbf{x})}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 F(\mathbf{x})}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 F(\mathbf{x})}{\partial x_2 \partial x_1} & \dfrac{\partial^2 F(\mathbf{x})}{\partial x_2^2} & \cdots & \dfrac{\partial^2 F(\mathbf{x})}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial^2 F(\mathbf{x})}{\partial x_n \partial x_1} & \dfrac{\partial^2 F(\mathbf{x})}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 F(\mathbf{x})}{\partial x_n^2} \end{bmatrix} \tag{35}$$

Hessian matricen er altid en symmetrisk matrix idet det gælder at

$$\frac{\partial^2 F(\mathbf{x})}{\partial x_i \partial x_j} = \frac{\partial^2 F(\mathbf{x})}{\partial x_j \partial x_i} \tag{36}$$

for alle differentiable funktioner (du vil ikke støde på ikke-differentiable funktioner i faget FR86).

## 5.2 Vektor differentiation af lineære former

En lineær form er en funktion af typen

$$F(\mathbf{x}) = \mathbf{a}'\mathbf{x} = \sum_{i=1}^{n} a_i x_i, \tag{37}$$

hvilket er det indre produkt af $\mathbf{x}$ med en vektor af konstante coefficienter $\mathbf{a}$. Da den partielt afledede af $F(\mathbf{x})$ med hensyn til $x_i$ er $a_i$, ses det umiddelbart at

$$\frac{\partial F(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{a}, \tag{38}$$

idet det $i$'te element i (38) er $a_i$.

## 5.3 Vektor differentiation af kvadratiske former

Lad $F(\mathbf{x})$ repræsentere den kvadratiske form af den symmetriske matrix $\mathbf{C}_{n\times n}$

$$F(\mathbf{x}) \;=\; \mathbf{x}'\mathbf{C}\mathbf{x} \;=\; \sum_{k=1}^{n}\sum_{l=1}^{n} x_k x_l c_{kl}. \tag{39}$$

Hvis vi differentierer (partielt) med hensyn til $x_i$ får vi

$$\begin{aligned} \frac{F(\mathbf{x})}{\partial x_i} &= \sum_{k=1}^{n}\sum_{l=1}^{n} \frac{\partial x_k}{\partial x_i} x_l c_{kl} + \sum_{k=1}^{n}\sum_{l=1}^{n} x_k \frac{\partial x_l}{\partial x_i} c_{kl} \\ &= \sum_{l=1}^{n} x_l c_{il} + \sum_{k=1}^{n} x_k c_{ki} \\ &= 2\sum_{k=1}^{n} c_{ik} x_k. \end{aligned} \tag{40}$$

Den første linie i (40) er den sædvanlige regneregel for differentiation af et produkt, mens forenklingen i den sidste linie skyldes at $\mathbf{C}$ er symmetrisk.

Hvis man kigger nærmere på den sidste linie i ligning (40), skulle man gerne genkende det som element $i$ fra multiplikationen mellem matricen $\mathbf{C}$ og vektoren $\mathbf{x}$. Dermed kan man, som generel regneregel for alle kvadratiske former, skrive

$$\frac{\partial \mathbf{x}'\mathbf{C}\mathbf{x}}{\partial \mathbf{x}} \;=\; 2\mathbf{C}\mathbf{x}. \tag{41}$$

Elementerne i Hessian matricen er givet ved

$$\frac{\partial^2 F(\mathbf{x})}{\partial x_i \partial x_j} \;=\; 2\sum_{k=1}^{n} c_{ik} \frac{\partial x_{ik}}{\partial x_j} \;=\; 2c_{ij}. \tag{42}$$

Det betyder med andre ord at hele Hessian matricen af den kvadratiske form kan skrives som

$$\frac{\partial^2 \mathbf{x}'\mathbf{C}\mathbf{x}}{\partial \mathbf{x}\partial \mathbf{x}'} \;=\; 2\mathbf{C}. \tag{43}$$

# 6 Et par anvendelser af matrix algebra

## 6.1 Stokastiske vektorer

Ved en stokastisk vektor $\mathbf{x}_n$ forstår vi en vektor af $n$ stokastiske variable $(x_1, x_2, \ldots, x_n)'$. Middelværdien af en stokastisk vektor er middelværdien af de enkelte elementer

$$E(\mathbf{x}) = \begin{bmatrix} E(x_1) \\ E(x_2) \\ \cdots \\ E(x_n) \end{bmatrix}. \tag{44}$$

Den flerdimensionale udgave af variansen er matricen

$$\begin{aligned} \mathrm{Cov}(\mathbf{x}) &= E\left[(\mathbf{x} - E(\mathbf{x}))(\mathbf{x} - E(\mathbf{x}))'\right] \\ &= \left[E\left\{x_i - E(x_i)\right\}\left\{x_j - E(x_j)\right\}\right]_{n\times n} \qquad (45) \\ &= \left[\mathrm{Cov}(x_i, x_j)\right]_{n\times n}. \qquad (46) \end{aligned}$$

Bemærk at det $ij$'te element i $n \times n$ matricen $\mathrm{Cov}(\mathbf{x})$ er kovariansen mellem de stokastiske variable $x_i$ og $x_j$. Da $\mathrm{Cov}(x_i, x_j) = \mathrm{Cov}(x_j, x_i)$, er matricen $\mathrm{Cov}(\mathbf{x})$ symmetrisk.

Man kan bruge ligning (45) til at vise at en kovariansmatrix altid som minimum er positiv semi-definit, og normalt endvidere altid positiv definit. Husk at en symmetrisk matrix $\mathbf{C}$ siges at være positiv definit hvis $\mathbf{w}'\mathbf{C}\mathbf{w} > 0$ for alle $\mathbf{w}$. Hvis vi lader $\mathrm{Cov}(\mathbf{x})$ være $\mathbf{C}$, kan vi skrive

$$\begin{aligned} \mathbf{w}'\mathrm{Cov}(\mathbf{x})\mathbf{w} &= \mathbf{w}'E\left[(\mathbf{x} - E(\mathbf{x}))(\mathbf{x} - E(\mathbf{x}))'\right]\mathbf{w} \\ &= E\left[\mathbf{w}'(\mathbf{x} - E(\mathbf{x}))(\mathbf{x} - E(\mathbf{x}))'\mathbf{w}\right] \qquad (47) \\ &= E\left[\left\{\mathbf{w}'(\mathbf{x} - E(\mathbf{x}))\right\}^2\right] > 0, \end{aligned}$$

idet den sidste linie er variansen på en skalar stokastisk variabel, og varianser er altid positive. Dermed har vi, bortset fra en enkelt undtagelse, vist at en generel kovariansmatrix altid er positiv definit.[2]

## 6.2 Variansen på en porteføljes afkast

Lad os antage at man kan investrer i $n$ aktiver, og at den stokastiske vektor $\mathbf{r}$ angiver afkastet i den næste periode for disse $n$ aktiver. Portefølje vægtene kalder vi $w_i$, og hvis vi samler dem i vektoren $\mathbf{w}$, kan porteføljens afkast skrives som

$$R_p = \sum_{i=1}^{n} w_i r_i = \mathbf{w}'\mathbf{r} \tag{48}$$

[2]Den pågældende undtagelse er hvis der eksisterer en linearkombination $\mathbf{w}$ af de $n$ stokastiske variable som altid er 0. Dette relativt sjældne tilfælde kaldes en stokastisk singularitet, og i så fald er kovariansmatricen kun positiv semi-definit, dvs. $\mathbf{w}'\mathbf{C}\mathbf{w} \geq 0$ for alle $\mathbf{w}$. Du vil faktisk støde på en variant af dette problem i forbindelse med estimation af de såkaldte heteroskedasticitet og autokorrelation konsistente kovariansmatricer, herunder Newey-West (1987) estimatoren.

Det forventede afkast på porteføljen er givet ved

$$E(R_p) = \sum_{i=1}^{n} w_i E(r_i) = \mathbf{w}' E(\mathbf{r}) \tag{49}$$

Variansen på porteføljens afkast kan beregnes som

$$\begin{aligned} \mathrm{Var}(R_p) &= E\left[\mathbf{w}'(\mathbf{r} - E(\mathbf{r}))\right]^2 \\ &= E\left[\mathbf{w}'(\mathbf{r} - E(\mathbf{r}))\left\{\mathbf{w}'(\mathbf{r} - E(\mathbf{r}))\right\}'\right] \\ &= E\left[\mathbf{w}'(\mathbf{r} - E(\mathbf{r}))(\mathbf{r} - E(\mathbf{r}))'\mathbf{w}\right] \\ &= \mathbf{w}' E\left[(\mathbf{r} - E(\mathbf{r}))(\mathbf{r} - E(\mathbf{r}))'\right]\mathbf{w} \\ &= \mathbf{w}'\mathrm{Cov}(\mathbf{r})\mathbf{w}. \end{aligned} \tag{50}$$

I den anden linie udnytter vi at transponeringen af en skalar også er en skalar, hvorved $x^2$ kan skrives som $xx'$ (en skalar kan opfattes som en $1 \times 1$ vektor). I den fjerde (næstsidste) linie flytter vi den konstante vektor $\mathbf{w}$ uden for forventningen, hvorefter vi får $\mathrm{Cov}(\mathbf{r})$. Bemærk at desuden at

$$\mathbf{w}'\mathrm{Cov}(\mathbf{r})\mathbf{w} = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_i \mathrm{Cov}(r_i, r_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_i \rho_{ij}\sigma_i\sigma_j, \tag{51}$$

hvor $\rho_{ij}$ er korrelationen mellem $r_i$ og $r_j$, og $\sigma_i^2$ er variansen på $r_i$.

## 6.3 Løsning af $n$ lineære ligninger med $n$ ubekendte

Matrix algebra er en effektiv metode til at løse $n$ lineære ligninger med $n$ ubekendte. Uden matrix algebra ville man opskrive ligningssystemet på følgende form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n. \end{aligned} \tag{52}$$

Hvis man i stedet samler elementerne $a_{ij}$ i en $n \times n$ matrix $\mathbf{A}$, og $x_i$ og $b_i$ i vektorerne $\mathbf{x}$ og $\mathbf{b}$, kan ligningssystemet skrives som

$$\mathbf{A}_{n\times n}\mathbf{x}_{n\times 1} = \mathbf{b}_{n\times 1}. \tag{53}$$

For at løse dette ligningssystem præmultiplicerer vi med $\mathbf{A}^{-1}$ på begge sider af lighedstegnet, hvilket giver

$$\mathbf{A}^{-1}\mathbf{A}\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} \;\Leftrightarrow\; \mathbf{x} = \mathbf{A}^{-1}\mathbf{b}, \tag{54}$$

idet $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ og $\mathbf{I}\mathbf{x} = \mathbf{x}$. Vi løser altså det linære ligningssystem ved først at invertere $\mathbf{A}$, og dernæst multiplicere $\mathbf{A}^{-1}$ og $\mathbf{b}$.

## 6.4 Multipel regression

Matrix algebra er desuden yderst anvendelig til multipel regression. I den klassiske regressions model, med $K-1$ forklarende variable (regressorer) foruden konstantledet, antager vi at den $i$'te observation $y_i$ kan skrives som

$$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \cdots + \beta_K x_{iK} + \varepsilon_i, \quad i = 1, 2, \ldots, n \tag{55}$$

hvor fejlleddet $\varepsilon_i$ opfylder betingelserne

$$\begin{aligned} E(\varepsilon_i) &= 0 & (56)\\ \text{var}(\varepsilon_i) &= \sigma^2 & (57)\\ \text{cov}(\varepsilon_i, \varepsilon_j) &= 0 \text{ for alle } i \neq j, & (58) \end{aligned}$$

hvilket betegnes som fravær at heteroskedasticitet og autokorrelation (hvis $y_i$ er en tidsserie).

Hvis vi nu definerer $\mathbf{x}_i = (1\ x_{i2}\ x_{i3}\ \cdots\ x_{iK})'$ og $\boldsymbol{\beta} = (\beta_1\ \beta_2\ \beta_3\ \cdots\ \beta_K)'$ som to $K$-dimensionale vektorer, kan vi skrive regressionsligningen (55) på følgende måde:

$$y_i = \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_i, \quad i = 1, 2, \ldots n \tag{59}$$

Det næste trin er er skrive regressionsligningen for samtlige $n$ observationer på matrix form:

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{\mathbf{y}_{n\times 1}} = \underbrace{\begin{bmatrix} \mathbf{x_1}' \\ \mathbf{x_2}' \\ \vdots \\ \mathbf{x_n}' \end{bmatrix} \boldsymbol{\beta}}_{\mathbf{X}_{n\times K}\boldsymbol{\beta}_{K\times 1}} + \underbrace{\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}}_{\boldsymbol{\varepsilon}_{n\times 1}}. \tag{60}$$

Ved hjælp af matrix algebra kan den fulde regressionsmodel altså skrives på meget kompakt form:

$$\mathbf{y}_{n\times 1} = \mathbf{X}_{n\times K}\boldsymbol{\beta}_{K\times 1} + \boldsymbol{\varepsilon}_{n\times 1}. \tag{61}$$

Bemærk at element $ij$ i matricen $\mathbf{X}$ er den $j$'te regressor (forklarende variabel) for den $i$'te observation. Vektoren er $\boldsymbol{\varepsilon}$ er en $n \times 1$ stokastisk vektor med egenskaberne

$$E(\boldsymbol{\varepsilon}) = 0_{n\times 1} \text{ og } \text{Cov}(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}_{n\times n}, \tag{62}$$

hvilket præcist er betingelserne (56)–(58) skrevet på matrixform.

Den mest almindelige estimator for den ukendte parameter(vektor) $\boldsymbol{\beta}$ er ordinary least squares (OLS), der på dansk kaldes mindste kvadraters metode. Estimatoren $\boldsymbol{\beta}$ findes ved at minimere funktionen

$$Q(\boldsymbol{\beta}) = \sum_{i=1}^{n} \varepsilon_i^2 = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}). \tag{63}$$

Ved at regne lidt på det sidste udtryk fås

$$Q(\boldsymbol{\beta}) = \mathbf{y}'\mathbf{y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}. \tag{64}$$

For at finde minimum for denne funktion skal vi beregne gradienten og sætte den lig med nul-vektoren. Da funktionen $Q(\boldsymbol{\beta})$ er en blanding af en lineær form og en kvadratisk form i $\boldsymbol{\beta}$, kan vi bruge regnereglerne fra afsnittet om vektor differentiation.

$$\frac{\partial Q(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\boldsymbol{\beta}. \tag{65}$$

At sætte gradienten lig med 0 svarer til at løse ligningssystemet

$$(\mathbf{X}'\mathbf{X})_{K\times K}\,\boldsymbol{\beta}_{K\times 1} = (\mathbf{X}'\mathbf{y})_{K\times 1}\,, \tag{66}$$

hvilket vi jo netop har gjort i afsnit 6.3. Løsningen er

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\,\mathbf{X}'\mathbf{y}, \tag{67}$$

hvor notationen $\hat{\boldsymbol{\beta}}$ betyder en estimator for den ukendte parameter $\boldsymbol{\beta}$. Det kan endvidere vises at kovariansmatricen for OLS estimatoren er givet ved

$$\mathrm{Cov}(\hat{\boldsymbol{\beta}}) = \sigma^2\,(\mathbf{X}'\mathbf{X})^{-1}\,. \tag{68}$$

I praksis erstatter man den ukendte $\sigma^2$ med en unbiased estimator, nemlig

$$s^2 = \frac{1}{n-K}\left(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right)'\left(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}\right), \tag{69}$$

som er summen af de kvadrerede residualer divideren med antallet af frihedsgrader.

# Literatur


Campbell, J.Y., A.W. Lo and A.C. MacKinlay (1997), *The Econometrics of Financial Markets*, Princeton University Press.

Newey, W. and K.D. West (1987), "A Simple Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.