Basic matrix methods and linear algebra

Lecture notes
MATLAB course at HiT

© Dr. ing.
David Di Ruscio

Telemark University College
Email: david.di.ruscio@hit.no
Porsgrunn, Norway

September 2000

Revised April 9, 2010

Porsgrunn **April 9, 2010**

Telemark University College
Kjølnes Ring 56
N-3914 Porsgrunn, Norway

Forord

This lecture notes contains basic theory in linear algebra and matrix methods.

Innhold

Kapittel 1

Introduction

Kapittel 2

Basic matrix and vector theory and computations

2.1 Vectors

An n-dimensional vector $x \ \ in \ (\mathbb{R})^n$ is a collection of n real numbers. These n real numbers can be organized in a array. We define a vector as a column of these figures.

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \in \ \mathbb{R}^n. \tag{2.1}$$

En n dimensional vector does also have one column with n rows.

The transpose of an n dimensional vector $x \in \mathbb{R}^n$ is denoted x^T and is a row-vector of the form

$$x^T = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \quad \in \ \mathbb{R}^{1\times n}. \tag{2.2}$$

Unless otherwise specified a vector is defined as a column vector as in (2.1).

Given two vectors with same dimension, ie $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^m$. The inner product is defined by

$$x^T y = y^T x = x_1 y_1 + x_2 y_2 + \cdots + x_m y_m = \sum_{k=1}^{m} x_k y_k \quad \in \ \mathbb{R} \tag{2.3}$$

Note that the inner product of two vectors is a scale. An important application of the inner product are to find the length of a vector. The length of the vector $x \in \mathbb{R}^n$ as defined in 2.1) is given by

$$||x|| = \sqrt{x^T x} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^n} \quad \in \ \mathbb{R} \tag{2.4}$$

A special case of this is Pythagoras statement that gives the length of a two-dimensional vector

$$c = \begin{bmatrix} a \\ b \end{bmatrix} \tag{2.5}$$

The length of the vector c is then

$$||c|| = \sqrt{c^T c} = \sqrt{a^2 + b^2} \tag{2.6}$$

We also note that a vector x is orthogonal to (perpendicular to) vector y if $x^T y = y^T x = 0$.

Example 2.1 *Given the vectors*

$$x = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}, \quad y = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} \tag{2.7}$$

Both vectors have the lebgth $||x|| = ||y|| = 3$ *because*

$$||x||^2 = 2^2 + 2^2 + (-1)^2 = 9 \tag{2.8}$$
$$||y||^2 = (-1)^2 + 2^2 + 2^2 = 9 \tag{2.9}$$

The vectors are orthogonal,i.e. the vectors are perpendicular to each other, because

$$x^T y = \begin{bmatrix} 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} = 2 \cdot (-1) + 2 \cdot 2 + (-1) \cdot 2 = 0 \tag{2.10}$$

2.2 Matriser

A matrix $A \in \mathbb{R}^{n \times m}$ is an array consisting of mn real numbers (figures).

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} \in \mathbb{R}^{n \times m}. \tag{2.11}$$

We remarks the following:

- An $n \times m$ real matrix A have n rows and m columns.
- When an uppercase letter denotes a matrix, e.g. as A as above, then the corresponding lowercase letter with sub-scripts ij, i.e., a_{ij}, is referred to as element (i, j) in the matrix A.

- the elements in a matrix may also be stored in a vector. This means that we can take all columns in a matrix and stack them above each other in a vector. This is usually the way matrices are stored in the computer. In fact we may only work with vectors in the computer and the software even if it virtually locks like we are working with matrices.
- The diagonal elements of the matrix A defines the vector

$$\text{diag(A)} = \begin{bmatrix} a_{11} \\ a_{22} \\ \vdots \\ a_{pp} \end{bmatrix} \in \mathbb{R}^{p=\min(n,m)}. \tag{2.12}$$

An example of a real 2×2 matrix is

$$A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \in \mathbb{R}^{2\times 2}. \tag{2.13}$$

The diagonal elements of the matrix A is then given by

$$\text{diag(A)} = \begin{bmatrix} 0 \\ -3 \end{bmatrix} \in \mathbb{R}^{2}. \tag{2.14}$$

2.3 The transpose of a matrix

The transpose of a matrix A is defined such that the matrix

$$C = A^T \tag{2.15}$$

have elements

$$c_{ij} = a_{ji} \tag{2.16}$$

An example of this is the matrix

$$A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \in \mathbb{R}^{2\times 2}. \tag{2.17}$$

where the transpose is given by

$$C = A^T = \begin{bmatrix} 0 & -2 \\ 1 & -3 \end{bmatrix} \in \mathbb{R}^{2\times 2}. \tag{2.18}$$

2.4 Special matrices

Innen lineær algebra har man definert mange spesielle matriser. Vi vil her nevne noen av de vanligste.

Identitetsmatrisen, I, opptrer svært ofte innen lineær algebra (matriseregning). En identitetsmatrise har enere på diagonalen og nuller utenom. mest vanlig er den kvardratiske identitetsmatrisen

$$I = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \quad \in \mathbb{R}^{n \times n}. \tag{2.19}$$

Man kan og tenke seg rektangulære identitetsmatriser, f. eks.

$$I_{2\times 3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \in \mathbb{R}^{2 \times 3}. \tag{2.20}$$

Ved rektangulære identitetsmatriser kan det være en fordel å oppgi dimensjonen, dvs. $I_{2\times 3}$ i dette tilfellet.

Diagonalmatriser er matriser som bare har elementer forskjellig fra null på hoved diagonalen. Et eksempel på en diagonalmatrise er

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \quad \in \mathbb{R}^{n \times n}. \tag{2.21}$$

Vi ser at identitetsmatrisen, I, er et spesialtilfelle av en diagonalmatrise.

Andre spesielle matriser er og som følger:

- Permutasjonsmatriser. En permutasjonsmatrise P er slik at den bytter om kolonner eller rekker i en matrise. En permutasjonsmatrise har egenskapene

$$P^T = P^{-1} \tag{2.22}$$

- Øvre triangulære matriser har bare elementer på og over diagonalen.
- Nedre triangulære matriser har bare elementer forskjellig fra null på og under diagonalen.
- **Ortogonale matriser**. En matrise $Q \in \mathbb{R}^{n \times m}$ er ortogonal dersom $Q^T Q = I$.

2.5 Matrisemultiplikasjon

Anta at vi har gitt to matriser $A \in \mathbb{R}^{n \times m}$ og $B \in \mathbb{R}^{m \times p}$. Matriseproduktet $C = AB$ er da gitt ved

$$C = AB \quad \in \mathbb{R}^{n \times p}. \tag{2.23}$$

Element c_{ij} i matrisen C er gitt ved

$$\begin{aligned} c_{ij} &= \sum_{k=1}^{m} a_{ik} b_{kj} \\ &= a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj} \end{aligned} \tag{2.24}$$

Dette kan gis følgende forklaringer:

- Element c_{ij} i matrisen $C = AB$ er gitt ved å ta produktsummen av alle elementene i rekke i i matrisen A med de korresponderende elementene i kolonne j i matrisen B.
- Element c_{ij} i matrisen $C = AB$ er gitt av indreproduktet av den i-te rekken i A med den j-te kolonnen i B.

Følgende dimensjonsanalyse ved matrisemultiplikasjon er meget nyttig.

$$n \left[\overset{m}{A} \right] m \left[\overset{p}{B} \right] = n \left[\overset{p}{C} \right] \tag{2.25}$$

Antall kolonner, m, i matrisen A må være lik antall rekker i matrisen B for at matriseproduktet $C = AB$ eksisterer. Vi ser at resultatet $C = AB$ blir en $n \times p$ matrise.

Man bør merke seg metoden for dimensjonsanalyse som vist i (2.25).

Legg merke til at for å implementere en matrisemultiplikasjon i et programmeringsspråk som f. eks. FORTRAN, C eller MATLAB så trenger vi tre for-løkker. En for-løkke for å implementere summen i (2.24), en for-løkke for å skanne over $i = 1, \ldots, n$ og en for-løkke for å skanne over $j = 1, \ldots, p$.

```
function C=dmamul(A,B)
% dmamul
% C=dmamul(A,B)
% Eksempel paa grunnleggende implementering av matrisemultiplikasjon vha.
% tre for-loekker.

[n,cols_a]=size(A);
[rows_b,p]=size(B);
if cols_a == rows_b
   m=cols_a;
   C=zeros(n,p);
   for i=1:n
      for j=1:p
         w=0;
         for k=1:m
            w=w+A(i,k)*B(k,j);
         end
         C(i,j)=w;
```

```
      end
   end
elseif cols_a ~= rows_b
   disp('Error:')
   C=NaN;
end
```

Gitt to matriser

$$A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \in \mathbb{R}^{2\times 2} \tag{2.26}$$

$$M = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \in \mathbb{R}^{2\times 2}. \tag{2.27}$$

Da er produktet AM gitt ved

$$AM = \begin{bmatrix} -1 & -2 \\ 1 & 4 \end{bmatrix}. \tag{2.28}$$

2.6 Addisjon av matriser og vektorer

Anta at vi har gitt to matriser $A \in \mathbb{R}^{n\times m}$ og $B \in \mathbb{R}^{n\times m}$ (NB: med samme dimensjon). Summen av matrisene er da gitt ved

$$C = A + B \quad \in \mathbb{R}^{n\times m}. \tag{2.29}$$

Det er viktig av man må ha samme dimensjon på matrisene eller vektorene man adderer. Summen $C = A + B$ vil dermed også ha samme dimensjon som A og B.

2.7 Determinanten til en matrise

Determinant er bare definert for kvadratiske matriser. Determinanten til en kvadratisk matrise $A \in \mathbb{R}^{n\times n}$ er definert ved den skalare "tallverdien"

$$\det(A) = |A| \tag{2.30}$$

Determinanten til en 2×2 matrise

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathbb{R}^{2\times 2}. \tag{2.31}$$

er gitt ved

$$\det(A) = |A| = a_{11}a_{22} - a_{21}a_{12}. \tag{2.32}$$

For vårt eksempel

$$A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \in \mathbb{R}^{2\times 2}. \tag{2.33}$$

er determinanten

$$\det(A) = |A| = 0 \cdot (-3) - (-2) \cdot 1 = 2 \tag{2.34}$$

Determinanten er også definert for høyere ordens matriser ($n > 2$). Vi viser her systemet for en 3×3 matrise

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \in \mathbb{R}^{3\times 3}. \tag{2.35}$$

Vi utvikler determinanten langs en rekke eller en kolonne. Her utvikler vi determinanten langs første kolonne

$$\det(A) = a_{11}\left|\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}\right| - a_{21}\left|\begin{bmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{bmatrix}\right| + a_{31}\left|\begin{bmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{bmatrix}\right| \tag{2.36}$$

Vi ser at determinanten til en høyere ordens matrise kan utrykkes som en sum av lavere ordens determinanter.

Vi skal spesielt merke oss at determinanten til en diagonal matrise er lik produktet av diagonalelementene, dvs. dersom

$$A = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} \in \mathbb{R}^{3\times 3}. \tag{2.37}$$

Da er

$$\det(A) = |A| = a_{11}a_{22}a_{33} \tag{2.38}$$

Det er og viktig å merke seg følgende i forbindelse med determinanter. Gitt to $n \times n$ kvadratiske matriser A og B. Da er

$$\det(AB) = \det(A)\det(B) \tag{2.39}$$
$$\det(A^T) = \det(A) \tag{2.40}$$
$$\det(cA) = c^n\det(A) \tag{2.41}$$
$$\det(A) \neq 0 \text{ når} A \text{ er inverterbar} \tag{2.42}$$
$$\det(A^{-1}) = \frac{1}{\det(A)} \tag{2.43}$$

Denne siste identiteten er enkel å vise fordi

$$AA^{-1} = I \tag{2.44}$$

som gir at

$$\det(A)\det(A^{-1}) = 1. \tag{2.45}$$

Ved å ta utgangspunkt i en egenverdidekomposisjon av matrisen A, dvs.

$$A = M\Lambda M^{-1} \tag{2.46}$$

ser vi at

$$\det(A) = \det(M\Lambda M^{-1}) = \det(\Lambda) = \prod_{i=1}^{n} \lambda_i \tag{2.47}$$

2.8 Invertering av matriser

Den inverse til en kvadratisk matrise $A \in \mathbb{R}^{n\times n}$ er den matrisen A^{-1} som (om den eksisterer) er slik at

$$AA^{-1} = A^{-1}A = I. \tag{2.48}$$

Det er spesielt viktig å merke seg at den inverse til en 2×2 matrise

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \in \quad \mathbb{R}^{2\times 2}. \tag{2.49}$$

er gitt ved

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \quad \in \quad \mathbb{R}^{2\times 2}. \tag{2.50}$$

For vårt eksempel

$$A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \tag{2.51}$$

har vi at

$$A^{-1} = \frac{1}{2} \begin{bmatrix} -3 & -1 \\ 2 & 0 \end{bmatrix} \tag{2.52}$$

Den inverse av en matrise A, dersom den eksisterer, kan generelt uttrykkes ved

$$A^{-1} = \frac{1}{\det A}(\mathrm{cof}(A))^T \tag{2.53}$$

Vi skal og merke oss at vi kan beregne den inverse av en matrise ved å løse ligningssystemet

$$AX = I \tag{2.54}$$

som gir at $X = A^{-1}$.

Vi tar et eksempel. Gitt

$$M = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -4 \\ 0 & 0 & 6 \end{bmatrix} \tag{2.55}$$

Da er Kofaktor matrisen, cof(A), blir:

$$\text{cof}A = \begin{bmatrix} 6 & 0 & 0 \\ +6 & 6 & 0 \\ 3 & +4 & 1 \end{bmatrix} \tag{2.56}$$

+ indikerer hvor vi har skiftet fortegn etter å ha satt opp matrisen med underdeterminanter. Videre har vi at

$$\det(A) = 6 \tag{2.57}$$

og at

$$A^{-1} = \frac{1}{6}\begin{bmatrix} 6 & 0 & 0 \\ +6 & 6 & 0 \\ 3 & +4 & 1 \end{bmatrix}^T = \begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 1 & \frac{2}{3} \\ 0 & 0 & \frac{1}{6} \end{bmatrix} \tag{2.58}$$

2.9 Egenverdier og egenvektorer

Tilhørende en kvardratisk matrise $A \in \mathbb{R}^{n\times n}$ kan vi definere følgende egenverdi- og egenvektor-problem

$$Am = \lambda m \tag{2.59}$$

der λ er en egenverdi og m en egenvektor til matrisen A. Dette betyr at

$$(\lambda I - A)m = 0 \tag{2.60}$$

Definisjon 2.1 (Egenverdier) *Egenverdiene til en matrise $A \in \mathbb{R}^{n\times n}$ er gitt ved de n røttene til den karakteristiske ligning, dvs.*

$$det(\lambda I - A) = 0. \tag{2.61}$$

Merk at polynomet $det(\lambda I - A)$ er definert som det karakteristiske polynom.

Vi har og, dersom egenvektormatrisen M er inverterbar,

$$A = M\Lambda M^{-1} \tag{2.62}$$

La oss se på følgende eksempelGitt matrisen

$$A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \tag{2.63}$$

Da finner vi at egenvetdiene er gitt ved

$$\lambda_1 = -1, \quad \lambda_2 = -2 \tag{2.64}$$

Egenvektormatrisen er da gitt ved

$$M = \begin{bmatrix} m_1 & m_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \tag{2.65}$$

2.10 Trace av en matrise

Tracen eller sporet til en kvadratisk matrise $A \in \mathbb{R}^{n\times n}$ er definert som summen av diagonalelementene i matrisen. Vi har

$$\begin{aligned} \text{trace}(A) &= \text{trace}(\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}) \\ &= a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii}. \end{aligned} \tag{2.66}$$

Anta at vi har gitt matrisedekomposisjonen

$$A = M\Lambda M^{-1} \tag{2.67}$$

Da er

$$\text{tr}(A) = \text{tr}(M\Lambda M^{-1}) = \text{tr}(M^{-1}M\Lambda) = \text{tr}(\Lambda) = \sum_{i=1}^{n} \lambda_i \tag{2.68}$$

2.11 Symmetriske matriser

I en symmetrisk matrise er alle elementene over diagonalen lik de tilsvarende elementer under diagonalen. Dvs. en matrise Q er symmetrisk dersom

$$q_{ij} = q_{ji} \tag{2.69}$$

En matrise

$$Q = \begin{bmatrix} a_{11} & a_{21} \\ a_{21} & a_{22} \end{bmatrix} \in \mathbb{R}^{2\times 2}. \tag{2.70}$$

er symmetrisk fordi $a_{21} = a_{12}$.

- Merk at alle egenverdiene i en symmetrisk og reell matrise vil være reelle.
- Symmetriske matriser har en egenvektormatrise som tilfredstiller $M^{-1} = M^T$. Dvs. slik at

$$Q = M\Lambda M^{-1} = M\Lambda M^T \tag{2.71}$$

2.12 Kvadratiske former og funksjoner

Gitt en symmetrisk matrise Q. Da definerer vi den skalare funksjonen

$$J = x^T Q x + 2f^T x + J_0 \tag{2.72}$$

for en kvadratisk funksjon. Spesielt så definerer vi

$$J = x^T Q x \tag{2.73}$$

for en **kvadratisk form**.

Kvadratiske funksjoner danner ofte konvekse funksjoner som har et unikt minimum. Vi finner minimum av funksjonen $J = x^T Q x + 2f^T x + J_0$ ved å sette den deriverte lik null, dvs.

$$\frac{\partial J}{\partial x} = 2Qx + 2f = 0 \tag{2.74}$$

som gir

$$x^* = -Q^{-1} f. \tag{2.75}$$

Vi har her benyttet derivasjonsregler for kvadratiske former som presentert i appendix.

2.13 Positiv definite matriser

Vi har og følgende definisjoner i forbindelse med symmetriske matriser:

- En matrise Q er positiv definit dersom

  $$J = x^T Q x > 0 \tag{2.76}$$

  for alle $x \neq 0$. Videre vil en positiv definit matrise ha bare positive og reelle egenverdier. Dette kan vi se ved å benytte egenverdidekomposisjonen $Q = M \Lambda M^T$.

- En matrise Q er positiv semi-definit dersom

  $$J = x^T Q x \geq 0 \tag{2.77}$$

  for alle $x \neq 0$. Dette betyr at egenverdiene til Q må være større eller lik null.

2.14 Singulære og ikke singulære matriser

- Vi sier at en matrise $A \in \mathbb{R}^{n \times n}$ er **singulær** dersom den ikke er inverterbar. En matrise er singulær dersom
  - $\det(A) = 0$
  - Dersom noen av egenverdiene til A er lik null.
- Vi sier at en matrise $A \in \mathbb{R}^{n \times n}$ er **ikke-singulær** dersom den inverterbar, dvs. dersom A^{-1} eksisterer. En matrise er ikke-singulær dersom
  - $\det(A) \neq 0$
  - Dersom ingen av egenverdiene til A er lik null.

2.15 Løsning av lineære ligninger og LU dekomposisjon

Vi skal her se på løsningen av et lineært ligningssystem. Gitt en kvardratisk matrise $A \in \mathbb{R}^{n\times n}$ og en vektor $b \in \mathbb{R}^n$. Vi definerer da det lineære ligningssystemet

$$Ax = b \tag{2.78}$$

der vektoren $x \in \mathbb{R}^n$ er ukjent. En åpenbar løsning er

$$x = A^{-1}b \tag{2.79}$$

dersom A er ikke singulær (inverterbar).

Dette er ingen effektiv måte å løse ligningsystemer på siden man først må beregne A^{-1} og deretter beregne produktet $A^{-1}b$.

En mer effektiv løsningsprosedyre er å benytte en eliminasjonsprosess, f.eks. **Gaus eliminasjon/transformasjon**. Dette går ut på å finne en serie transformasjoner M_1, ..., M_{n-1} slik at

$$M_{n-1}\dots M_2M_1A = U \tag{2.80}$$

er øvre triangulær. Vi kan dermed løse det ekvivalente ligningssystemet

$$Ux = M_{n-1}\dots M_2M_1b \tag{2.81}$$

med det vi kaller "back-substitution".

Gitt

$$A = \begin{bmatrix} 1 & 4 \\ 2 & 5 \end{bmatrix},\ b = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{2.82}$$

Vi har da at

$$M_1A = \begin{bmatrix} 1 & 4 \\ 0 & -3 \end{bmatrix},\ M_1b = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \tag{2.83}$$

der

$$M_1 = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}, \tag{2.84}$$

er nedre triangulær. Vi løser så ligningsystemet $Ux = M_1b$ mht. x, dvs.

$$\begin{bmatrix} 1 & 4 \\ 0 & -3 \end{bmatrix}, \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \tag{2.85}$$

som gir

$$x_2 = \frac{1}{3}, \tag{2.86}$$

$$x_1 = 1 - 4x_2 = -\frac{1}{3}. \tag{2.87}$$

LU-dekomposisjon er en systematisk form for Gaus-eliminasjon som bla. benyttes til løsning av lineære ligningssystemer. LU-dekomposisjon er en "høynivåbeskrivelse av Gaus-eliminasjon. Den er definert som følger

Definisjon 2.2 (LU-dekomposisjon)
Gitt en matrise $A \in \mathbb{R}^{n\times m}$. Da eksisterer det en nedre triangulær matrise $L \in \mathbb{R}^{n\times n}$ og en øvre triangulær matrise $U \in \mathbb{R}^{m\times m}$ slik at

$$A = LU. \tag{2.88}$$

Merk: vi antar at A er "singulær".

Dette kan benyttes til å løse $Ax = b$ ved først å løse

$$Ly = b \tag{2.89}$$

mht. y og deretter løse

$$Ux = y \tag{2.90}$$

mht. x ved "back-substitution".

La oss til slutt se på LU-dekomposisjon i forbindelse med vårt standardeksempel

$$A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \tag{2.91}$$

Da har vi at $A = LU$ der

$$L = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\ U = \begin{bmatrix} -2 & -3 \\ 0 & 1 \end{bmatrix} \tag{2.92}$$

2.16 Matrisenormer

Størrelsen av en matrise kan måles vha. begrepet matrisenorm. Noen av de viktigste matrisenormene er:

Frobeniusnormen

- Frobeniusnorm er definert ved

$$||A||_F^2 = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}^2 \tag{2.93}$$

  Frobeniusnormen er relatert til trace-begrepet via.

$$||A||_F^2 = \mathrm{tr}(A^T A) \tag{2.94}$$

- En viktig egenskap ved Frobeniusnormen er at den er invariant for ortogonale transformasjoner. Dvs. for alle ortogonale matriser Q og U med passende dimensjon har vi at

$$||QAU||_F = ||A||_F \tag{2.95}$$

2-normen til en matrise A er definert ved

$$||A||_2^2 = \lambda_{\max}(A^T A) = \sigma_{\max}(A) \tag{2.96}$$

der $\sigma_{\max}(A)$ er den største singulærverdien til matrisen A.

La oss som et eksempel studere

$$A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \tag{2.97}$$

Vi har da at

$$||A||_F = \sqrt{1+4+9} = \sqrt{14} = 3.7417 \tag{2.98}$$

2-normen til A er gitt ved egenverdiene til

$$A^T A = \begin{bmatrix} 4 & 6 \\ 6 & 10 \end{bmatrix} \tag{2.99}$$

som er $\lambda_1 = 0.2918$, $\lambda_2 = 13.7082$. dette gir at

$$||A||_2 = \sqrt{\lambda_{\max}(A^T A)} = \sqrt{13.7082} = 3.7025. \tag{2.100}$$

2.17 Minste kvadraters metode

Gitt et lineært overbestemt ligningssystem

$$Y = XB + E \tag{2.101}$$

der $Y \in \mathbb{R}^N$, $X \in \mathbb{R}^{N \times r}$ kjente datamatriser. E er en støymatrise. $B \in \mathbb{R}^r$ er en ukjent vektor av regressjonsparametre.

Siden ligningsystemet er overbestemt så finnes det mange løsninger X. Anta at vi ønsker å finne den løsning B_{OLS} slik at den kvadratiske funksjonen

$$V = (Y - XB)^T (Y - XB) = ||Y - XB||_F^2 \tag{2.102}$$

minimaliseres. OLS løsningen er gitt ved

$$B_{OLS} = (X^T X)^{-1} X^T Y \tag{2.103}$$

dersom $X^T X$ er ikke-singulær. Dette er ekvivalent med at X må ha rang m.

Den optimale OLS prediksjonen av Y er da gitt ved

$$Y_{OLS} = XB_{OLS} = X(X^T X)^{-1} X^T Y \tag{2.104}$$

La oss se på et eksempel der

$$X = \begin{bmatrix} 0 & -1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix}, \; Y = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \tag{2.105}$$

Minste kvadraters metode løsningen er da gitt ved

$$B_{OLS} = (X^T X)^{-1} X^T Y = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} \tag{2.106}$$

Vi har benyttet at

$$X^T X = \begin{bmatrix} 2 & -2 \\ -2 & 3 \end{bmatrix}, \; X^T Y = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \tag{2.107}$$

2.18 Ortogonale projeksjoner

- A matrix Y can be decomposed into two matrices with orthogonal row spaces.

$$Y = Y/P + YP^{\perp}$$

- Projection of the row space of Y onto the row space of P.

$$Y/P = YP^T(PP^T)^{\dagger}P \tag{2.108}$$

- Projection of the row space of Y onto the orthogonal complement of the row space of P.

$$YP^{\perp} = Y - YP^T(PP^T)^{\dagger}P \tag{2.109}$$

Some useful results

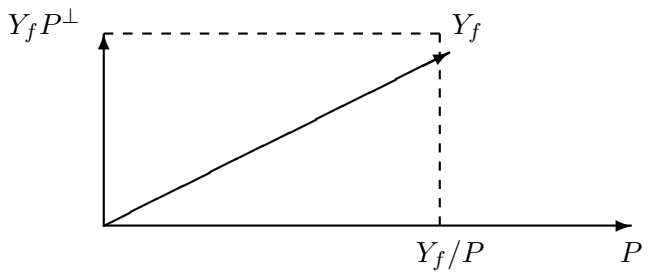
Lemma 2.18.1 *The following equality is true*

$$U/\begin{bmatrix} U \\ W \end{bmatrix} = U \tag{2.110}$$

Lemma 2.18.2 *The following equality is true*

$$U\begin{bmatrix} U \\ W \end{bmatrix}^{\perp} = 0 \tag{2.111}$$

La oss i forbindelse med dette studere to problemer.



Figur 2.1: Two dimensional illustration of orthogonal projections.

Minste kvadraters metode og projeksjoner

Gitt minste kvadraters metode problemet

$$Y^T = B^T X^T + E^T \tag{2.112}$$

dvs

$$\mathcal{Y} = \mathcal{B}\mathcal{X} + \mathcal{E} \tag{2.113}$$

Da har vi at

$$\mathcal{Y}_{OLS} = \mathcal{B}_{OLS}\mathcal{X} = \mathcal{Y}/\mathcal{X} \tag{2.114}$$

Systemorden

Vi venter med dette til vi har presentert SVD.

2.19 QR-dekomposisjon

En matrise $A \in \mathbb{R}^{N\times m}$ der $N \geq m$ kan faktoriseres slik at

$$A = QR \tag{2.115}$$

der $R \in \mathbb{R}^{N\times m}$ er en øvre triangulær matrise og $Q \in \mathbb{R}^{N\times N}$ er en orthogonal matrise slik at $Q^T = Q^{-1}$.

- QR-dekomposisjonen benyttes i flere algoritmer for underromsbasert (subspace) systemidentifikasjon, f. eks. i DSR.
- QR-dekomposisjonen benyttes ved løsning av minste kvadraters problemer og lineære ligningsystemer.

La oss se på løsning av

$$Y = XB + E \tag{2.116}$$

ved hjelp av QR dekomposisjon. Her er $Y \in \mathbb{R}^N$, $X \in \mathbb{R}^{N \times r}$ kjente datamatriser. E er en støymatrise. $B \in \mathbb{R}^r$ er en ukjent vektor av regressjonsparametre. QR-dekomposisjon gir at

$$X = QR \tag{2.117}$$

Vi har da at

$$Q^T Y = RB \tag{2.118}$$

som kan løses enkelt. Legg merke til at

$$||XB - Y||_F^2 = ||Q^T XB - Q^T Y||_F^2 = ||R_1 B - c_1||_F^2 \tag{2.119}$$

der R_1 er den øvre kvadratiske delen av R og c_1 den tilsvarende øvre delen av $c = Q^T R$.

La oss se på et eksempel der

$$X = \begin{bmatrix} 0 & -1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix}, \; Y = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \tag{2.120}$$

En QR-dekomposisjon av X gir

$$R = \begin{bmatrix} -1.4142 & 1.4142 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \; Q = \begin{bmatrix} 0 & -1 & 0 \\ 0.7071 & -0.0000 & 0.7071 \\ -0.7071 & 0.0000 & 0.7071 \end{bmatrix} \tag{2.121}$$

Vi står da igjen med ligningsystemet

$$c_1 = R_1 B \tag{2.122}$$

der

$$R_1 = \begin{bmatrix} -1.4142 & 1.4142 \\ 0 & 1 \end{bmatrix}, \; c_1 = \begin{bmatrix} 0.7071 \\ 1 \end{bmatrix} \tag{2.123}$$

fordi

$$Q^T Y = c = \begin{bmatrix} 0.7071 \\ 1 \\ 0.7071 \end{bmatrix} \tag{2.124}$$

Vi finner at løsningen er

$$B_{OLS} = (X^T X)^{-1} X^T Y = R_1^{-1} c_1 = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} \tag{2.125}$$

2.20 Data komprimering og QR dekomposisjon

Define the following standard QR decomposition

$$\frac{1}{\sqrt{N}}\tilde{Y} = \frac{1}{\sqrt{N}}\begin{bmatrix} X^T \\ Y^T \end{bmatrix} = \begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix}\begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} = RQ \tag{2.126}$$

where

$$R_{11} \in \Re^{r\times r} \quad R_{21} \in \Re^{m\times r} \quad R_{22} \in \Re^{m\times m} \tag{2.127}$$

$$R \in \Re^{(r+m)\times(r+m)} \quad Q \in \Re^{(r+m)\times N} \tag{2.128}$$

The solution to the total multivariate problem is given by the triangular factors R_{11}, R_{21} and R_{22}, only. The orthogonal matrix Q is not needed. This will reduce the computational effort and storage considerably, especially when the number of observations N is large compared to the number of variables.

We have directly the following equation for the regression coefficients

$$R_{21} = \mathbf{B}^T R_{11} \tag{2.129}$$

In order to solve this equation for $\mathbf{B}$, standard PLS or PCR methods can be applied.

The lower triangular matrix R_{22} is the square root of the residual covariance matrix. The covariance estimate of the noise (or residuals) is given by

$$\hat{\Delta} = \frac{1}{N}E^T E = R_{22}R_{22}^T \tag{2.130}$$

2.21 Andre matrisedekomposisjoner

Man har følgende viktige matrisedekomposisjoner innen lineær algebra:

- Cholesky faktorisering/dekomposisjon eller kvadratrotsfaktorisering av symmetriske matriser.

  $$Q = RR^T \tag{2.131}$$

  der R er en øvre triangulær matrise. Den kvadratiske formen

  $$J = x^T Qx \tag{2.132}$$

  kan da uttrykkes ved indreproduktet

  $$J = y^T y \tag{2.133}$$

  der

  $$y = R^T x. \tag{2.134}$$

  Cholesky faktorisering kalles også i mange sammenhenger for kvadratrotfaktorisering og benyttes f. eks. i Biermans effektive implementering av Kalman-filteret.

- Singulærverdidekomposisjon (SVD). Dette er en meget viktig dekomposisjon som benyttes bla. til å beregne rangen til en matrise.

- QR-dekomposisjon. En matrise kan faktoriseres slik:

$$A = QR \tag{2.135}$$

der R er en øvre triangulær matrise og Q er en orthogonal matrise slik at $Q^T = Q^{-1}$. QR-dekomposisjonen benyttes i flere algoritmer for subspace systemidentifikasjon, f. eks. i DSR.

- Schur dekomposisjon. En kvadratisk matrise A kan dekomponeres slik at

$$A = UTU^T \tag{2.136}$$

der T er en øvre triangulær matrise med egenverdiene til A på diagonalen. Dersom A har komplekse egenverdier så vil disse finnes som egenverdiene til 2×2 blokker på diagonalen til T. U er en ortogonal matrise slik at $U^{-1} = U^T$ og $U^TU = I$.

En viktig egenskap med Schur dekomposisjonen er at den alltid eksisterer selv om A har multiple egenverdier.

La oss se på en Schur dekomposisjon til

$$A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \tag{2.137}$$

som gir

$$T = \begin{bmatrix} -1 & 3 \\ 0 & -2 \end{bmatrix}, \; U = \begin{bmatrix} 0.7071 & 0.7071 \\ -0.7071 & 0.7071 \end{bmatrix}, \tag{2.138}$$

2.22 The Singular Value Decomposition

Let A be an $n \times m$ real matrix. The Singular value Decomposition (SVD) of the matrix A is then defined as

$$A = USV^T = \sum_{i=1}^{p} s_i u_i v_i^T, \tag{2.139}$$

where $U \in \mathbb{R}^{n \times n}$ is an orthogonal matrix of left-hand-side singular vectors and $V \in \mathbb{R}^{m \times m}$ is an orthogonal matrix of right-hand-side singular vectors, i.e.,

$$U = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \tag{2.140}$$

$$V = \begin{bmatrix} v_1 & v_2 & \cdots & v_m \end{bmatrix} \tag{2.141}$$

$u_i \in \mathbb{R}^n \; \forall \; i = 1, \ldots, n$ is defined as the left-hand-side singular vectors and $v_i \in \mathbb{R}^m \; \forall \; i = 1, \ldots, m$ is defined as the right-hand-side singular vectors. Furthermore, since U and V are orthogonal matrices we have that

$$U^TU = UU^T = I_n, \tag{2.142}$$

and

$$VV^T = V^TV = I_m, \tag{2.143}$$

where I_n and I_m are the $n \times n$ and the $m \times m$ identity matrices, respectively.

$S \in \mathbb{R}^{n\times m}$ is a diagonal matrix of singular values $s_i \ \forall \ i = 1, \ldots, p$ where the number of singular values are $p = \min(n, m)$. Furthermore, the singular values are positive scalar numbers such that

$$s_1 \geq s_2 \geq \cdots \geq s_p \geq 0. \tag{2.144}$$

Example 2.2 *Consider the matrix*

$$A = \begin{bmatrix} 0.96 & 1.72 \\ 2.28 & 0.96 \end{bmatrix}. \tag{2.145}$$

The SVD of A is then given by

$$A = USV^T = \overbrace{\begin{bmatrix} 0.6 & -0.8 \\ 0.8 & 0.6 \end{bmatrix}}^{U} \overbrace{\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}}^{S} \overbrace{\begin{bmatrix} 0.8 & 0.6 \\ 0.6 & -0.8 \end{bmatrix}^T}^{V^T}. \tag{2.146}$$

Hence,

$$U = \begin{bmatrix} u_1 & u_2 \end{bmatrix} = \begin{bmatrix} 0.6 & -0.8 \\ 0.8 & 0.6 \end{bmatrix}, \tag{2.147}$$

$$V = \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} 0.8 & 0.6 \\ 0.6 & -0.8 \end{bmatrix}, \tag{2.148}$$

$$S = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}. \tag{2.149}$$

Furthermore, since the singular values, $s_1 = 3$ and $s_2 = 1$, are non-zero, we have that

$$rank(A) = 2 \tag{2.150}$$

Example 2.3 *Consider the Hankel matrix*

$$Y_{0|3} = \begin{bmatrix} 0 & 1 & 1.5 \\ 1 & 1.5 & 1.55 \\ 1.5 & 1.55 & 1.275 \end{bmatrix} \tag{2.151}$$

the SVD of $Y_{0|3}$ is then given by

$$Y_{0|3} = USV^T, \tag{2.152}$$

where,

$$U = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} = \begin{bmatrix} 0.4257 & 0.8293 & 0.3620 \\ 0.6310 & 0.0146 & -0.7756 \\ 0.6485 & -0.5586 & 0.5171 \end{bmatrix} \tag{2.153}$$

$$S = \begin{bmatrix} s_1 & 0 & 0 \\ 0 & s_2 & 0 \\ 0 & 0 & s_3 \end{bmatrix} = \begin{bmatrix} 3.7677 & 0 & 0 \\ 0 & 0.9927 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{2.154}$$

$$V = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} 0.4257 & -0.8293 & 0.3620 \\ 0.6310 & -0.0146 & -0.7756 \\ 0.6485 & 0.5586 & 0.5171 \end{bmatrix} \tag{2.155}$$

Since the last singular value, $s_3 = 0$, we conclude that

$$rank(Y_{0|3}) = 2 \tag{2.156}$$

Hence, the SVD of $Y_{0|3}$ can also be written as

$$Y_{0|3} = U_1 S_1 V_1^T, \tag{2.157}$$

where

2.23 Rang av matriser

- Rangen til en matrise A er lik antall singulærverdier som er forskjellig fra null.
- $$\text{rank}(A) = \text{rank}(A^T) \tag{2.158}$$
- $$\text{rank}(AB) = \text{rank}(A) \tag{2.159}$$

  dersom B har full rang.
- $$\text{kolonne rangen} = \text{rekke rangen} \tag{2.160}$$
- Sylvesters ulikhet

$$\text{rank}(A) + \text{rank}(B) \leq \text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B)) \tag{2.161}$$

Gitt en matriseligning

$$Y = OX + HU \tag{2.162}$$

der Y og U er kjente rektangulære matriser. Sett at vi ønsker å estimere O samt å finne rangen til matrisen O. Vi forutsetter at X har full rang n. Vi har at

$$Z = YU^{\perp} = OXU^{\perp} \tag{2.163}$$

Dette gir at

$$\text{rank}(O) = \text{rank}(YU^{\perp}) \tag{2.164}$$

Videre så kan vi estimere O ut i fra en SVD av Z.

2.24 Kondisjonstallet til en matrise

Kondisjonstallet til en matrise sier noe om inverterbarheten av en matrise.

$$\text{cond}(A) = \frac{\sigma_1}{\sigma_p} \tag{2.165}$$

der σ_1 er den største og σ_p den minste singulærverdien til matrisen A.

Dersom kondisjonstallet til en matrise A er $\text{cond}(A) = \infty$ så er matrisen singulær, dvs. ikke inverterbar.

2.25 Vektor operator og kronecker produkt

Elementene i en matrise kan også lagres i en vektor. Dvs. man kan ta alle kolonnene i matrisen og stable oppå hverandre i en vektor. Det er slik man fysisk lagrer matriser i en datamaskin.

Vi definerer vektoroperatoren

$$\text{vec}(A) = \begin{bmatrix} a_{11} \\ \vdots \\ a_{n1} \\ a_{12} \\ \vdots \\ a_{n2} \\ \vdots \\ a_{1m} \\ \vdots \\ a_{nm} \end{bmatrix} \in \mathbb{R}^{nm}. \tag{2.166}$$

Anta at vi har gitt en lineær matrise ligning

$$Y = XB + E \tag{2.167}$$

der $Y \in \mathbb{R}^{N\times m}$, $X \in \mathbb{R}^{N\times r}$ kjente datamatriser. E er en støymatrise. $B \in \mathbb{R}^{r\times m}$ er en ukjent vektor av regressjonsparametre. Denne matriseligningen kan uttrykkes som vektorligningen

$$\text{vec}(Y) = (I_m \otimes X)\text{vec}(B) + \text{vec}(E) \tag{2.168}$$

der $\text{vec}(Y) \in \mathbb{R}^{Nm}$ er en kolonne vektor, $(I_m \otimes X) \in \mathbb{R}^{Nm\times rm}$ and $\text{vec}(B) \in \mathbb{R}^{rm}$.

2.26 PCA og PCR

Anta at vi har gitt en lineær matrise ligning

$$Y = XB + E \tag{2.169}$$

der X kan ha multikolinneære kolonner.

2.27 Cayley hamiltons teorem

Se appendix B.

2.28 PLS

Se artikkel.

2.29 Oppgaver

a) Hva er et indreprodukt ?

b) Hvordan beregner man lengden til en vektor ?

c) Hva menes med et yttreprodukt ?

d) Hva menes med en identitetsmatrise ?

e) Hva menes med en diagonalmatrise ?

f)

Appendix A

More about linear algebra and matrix methods

A.1 Trace of a matrix

The trace of a $n \times m$ matrix A is defined as the sum of the diagonal elements of the matrix, i.e.

$$tr(A) = \sum_{i=1}^{n} a_{ii} \tag{A.1}$$

We have the following trace operations on two matrices A and B of apropriate dimensions

$$tr(A^T) = tr(A) \tag{A.2}$$
$$tr(AB^T) = tr(A^T B) = tr(B^T A) = tr(BA^T) \tag{A.3}$$
$$tr(AB) = tr(BA) = tr(B^T A^T) = tr(A^T B^T) \tag{A.4}$$
$$tr(A \pm B) = tr(A) \pm tr(B) \tag{A.5}$$

A.2 Gradient matrices

$$\frac{\partial}{\partial X} \quad tr[X] = I \tag{A.6}$$
$$\frac{\partial}{\partial X} \quad tr[AX] = A^T \tag{A.7}$$
$$\frac{\partial}{\partial X} \quad tr[AX^T] = A \tag{A.8}$$
$$\frac{\partial}{\partial X} \quad tr[AXB] = A^T B^T \tag{A.9}$$
$$\frac{\partial}{\partial X} \quad tr[AX^T B] = BA \tag{A.10}$$
$$\frac{\partial}{\partial X} \quad tr[XX] = 2X^T \tag{A.11}$$
$$\frac{\partial}{\partial X} \quad tr[XX^T] = 2X \tag{A.12}$$

$$\frac{\partial}{\partial X} \quad tr[X^n] = n(X^{n-1})^T \tag{A.13}$$
$$\frac{\partial}{\partial X} \quad tr[AXBX] = A^T X^T B^T + B^T X^T A^T \tag{A.14}$$
$$\frac{\partial}{\partial X} \quad tr[e^{AXB}] = (Be^{AXB}A)^T \tag{A.15}$$

$$\frac{\partial}{\partial X} tr[XAX^T] = 2XA, \text{ if } A = A^T \tag{A.16}$$

$$\frac{\partial}{\partial X^T} \quad tr[AX] = A \tag{A.17}$$
$$\frac{\partial}{\partial X^T} \quad tr[AX^T] = A^T \tag{A.18}$$
$$\frac{\partial}{\partial X^T} \quad tr[AXB] = BA \tag{A.19}$$
$$\frac{\partial}{\partial X^T} \quad tr[AX^T B] = A^T B^T \tag{A.20}$$
$$\frac{\partial}{\partial X^T} \quad tr[e^{AXB}] = Be^{AXB}A \tag{A.21}$$

A.3 Derivatives of vector and quadratic form

The derivative of a vector with respect to a vector is a matrix. We have the following identities:

$$\frac{\partial x}{\partial x^T} \quad = I \tag{A.22}$$
$$\frac{\partial}{\partial x} \quad (x^T Q) = Q \tag{A.23}$$
$$\frac{\partial}{\partial x} \quad (Qx) = Q^T \tag{A.24}$$
$$\tag{A.25}$$

The derivative of a scalar with respect to a vector is a vector. We have the following identities:

$$\frac{\partial}{\partial x} \quad (y^T x) = y \tag{A.26}$$
$$\frac{\partial}{\partial x} \quad (x^T x) = 2x \tag{A.27}$$
$$\frac{\partial}{\partial x} \quad (x^T Qx) = Qx + Q^T x \tag{A.28}$$
$$\frac{\partial}{\partial x} \quad (y^T Qx) = Q^T y \tag{A.29}$$

Note that if Q is symmetric then

$$\frac{\partial}{\partial x} \quad (x^T Qx) = Qx + Q^T x = 2Qx. \tag{A.30}$$

A.4 Matrix norms

The trace of the matrix product $A^T A$ is related to the Frobenius norm of A as follows

$$\|A\|_F^2 = \text{tr}(A^T A), \tag{A.31}$$

where $A \in \mathbb{R}^{N \times m}$.

A.5 Linearization

Given a vector function $f(x) \in \mathbb{R}^m$ where $x \in \mathbb{R}^n$. The derivative of the vector f with respect to the row vector x^T is defined as

$$\frac{\partial f}{\partial x^T} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \dots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{m \times n} \tag{A.32}$$

Given a non-linear differentiable state space model

$$\dot{x} = f(x, u), \tag{A.33}$$
$$y = g(x). \tag{A.34}$$

A linearized model around the stationary points x_0 and u_0 is

$$\dot{\delta x} = Ax + Bu, \tag{A.35}$$
$$\delta y = Dx, \tag{A.36}$$

where

$$A = \frac{\partial f}{\partial x^T} \mid_{x_0, u_0}, \tag{A.37}$$
$$B = \frac{\partial f}{\partial u^T} \mid_{x_0, u_0}, \tag{A.38}$$
$$D = \frac{\partial g}{\partial x^T} \mid_{x_0, u_0}, \tag{A.39}$$

and where

$$x = x - x_0, \tag{A.40}$$
$$u = u - u_0. \tag{A.41}$$

A.6 Kronecer product matrices

Given a matrix $X \in \mathbb{R}^{N \times r}$. Let I_m be the $(m \times m)$ identity matrix. Then

$$(X \otimes I_m)^T = X^T \otimes I_m, \tag{A.42}$$

$$(I_m \otimes X)^T = I_m \otimes X^T. \tag{A.43}$$

Appendix B

Basic system theory

B.1 Models of dynamic systems

The aim of this section is not to discuss modeling principles of dynamic systems in detail. However we will in this introductory section mention that dynamic models may be developed in many ways. For instance so called first principles methods as mass balances, force balances, energy balances, i.e., conservation of law methods, leads to ether non-linear models of the type

$$\dot{x} = f(x,u) \tag{B.1}$$
$$y = g(x) \tag{B.2}$$

or linear or linearized models of the type

$$\dot{x} = Ax + Bu \tag{B.3}$$
$$y = Dx \tag{B.4}$$

Note also that a linearized approximation of the non-linear model usually exist. We will in the following give a simple example of a system which may be described by a linear continuous time state space model

Example B.1 (Model of a damped spring system)
Assume given an object with mass, m, influenced by three forces. One force F_1 used to pull the mass, one force $F_2 = kx$ from the spring and one force $F_3 = \mu\dot{x} = \mu v$ that represents the friction or viscous damping.

We define x as the position of the object and $\dot{x} = v$ as the velocity of the object. Furthermore the force F_1 may be defined as a manipulable control input variable and we use u as a symbol for this control input, i.e., $u = F_1$.

from this we have the following force balance

$$ma = m\dot{v} = \sum_{i=1}^{3} F_i = F_1 - F_2 - F_3 = -kx - \mu v + u \tag{B.5}$$

The model for the damped spring system consists of two continuous time ordinary differential equations. Those two ODEs may be written in standard state space form as follows

$$\overbrace{\begin{bmatrix} \dot{x} \\ \dot{v} \end{bmatrix}}^{\dot{x}} = \overbrace{\begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{\mu}{m} \end{bmatrix}}^{A} \overbrace{\begin{bmatrix} x \\ v \end{bmatrix}}^{x} + \overbrace{\begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix}}^{B} u \tag{B.6}$$

Modeling from first principles, e.g., as the in the damped spring example above, often leads to a standard linear continuous time state space model on the form

$$\dot{x} = Ax + Bu \tag{B.7}$$

where $x \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}^r$ is the control input vector, $A \in \mathbb{R}^{n\ timesn}$ is state matrix and $B \in \mathbb{R}^{n\ timesr}$ is the control input matrix.

B.2 State space Models

An important class of state space models is the time invariant linear and continuous state space model of the form

$$\dot{x} = Ax + Bu, \qquad x(0) = x_0, \tag{B.8}$$

$$y = Dx, \tag{B.9}$$

where $u \in \mathbb{R}^r$ is the control vector, $x \in \mathbb{R}^n$ is the state vector, $y \in \mathbb{R}^m$ is the measurements vector and $x_0 = x(t_0) \in \mathbb{R}^n$ is the initial value of the state vector, which usually is assumed to be known.

It can be shown that the exact solution of the state equation (B.8) at time $t_0 \leq t$ is given by

$$x(t) = e^{A(t-t_0)}x(t_0) + \int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)d\tau. \tag{B.10}$$

As we see, the solution consists of two parts. The first part represents the autonomous response (homogenous solution) driven only by initial values different from zero. The second term represents the in homogenous solution driven by the control variable, $u(t)$.

In order to compute the first term we have to compute the matrix exponential $e^{A(t-t_0)}$. This matrix exponential is defined as the transition matrix, because it defines the transition of the state from the initial value, $x(t_0)$, to the final state $x(t)$ in an autonomous system $\dot{x} = Ax$ with known initial state $x(t_0)$. The transition matrix is defined as follows

$$\Phi(t) \stackrel{\text{def}}{=} e^{At}. \tag{B.11}$$

Using this definition of the transition matrix we see that the solution (B.10) can be written as follows

$$x(t) = \Phi(t - t_0)x(t_0) + \int_{t_0}^{t} \Phi(t - \tau)Bu(\tau)d\tau. \tag{B.12}$$

The second term in the solution (B.10) (ore equivalent as in (B.12)) consists of a convolutional integral. This integral must usually be computed numerically, e.g. it is usually hard to obtain an analytically solution. However, an important special case is the case where the control $u(\tau)$ is constant over the integration interval $t_0 < \tau \leq t$.

$$x(t) = \Phi(t - t_0)x(t_0) + \Delta u(t_0), \tag{B.13}$$

where Δ is shown to be

$$\Delta = \int_{t_0}^{t} e^{A(t-\tau)}Bd\tau = \int_{0}^{t-t_0} e^{A\tau}Bd\tau \tag{B.14}$$

Note also that

$$\Delta = A^{-1}(e^{A(t-t_0)} - I)B, \tag{B.15}$$

when A is non singular. It is this solution which usually is used in order to compute the general solution to the state equation. Hence, the control input $u(t)$ is assumed to be constant over piece wise identical intervals $\Delta t = t - t_0$.

The constant interval Δt is in control theory and control systems defined as the sampling time in the digital controller. If we now are putting $t = t_0 + \Delta t$ in the solution (B.13) then we get

$$x(t_0 + \Delta t) = \Phi(\Delta t)x(t_0) + \Delta u(t_0), \tag{B.16}$$

where Δ is given by

$$\Delta = A^{-1}(e^{A\Delta t} - I)B. \tag{B.17}$$

The solution given by (B.16) and (B.17) is the starting point for making a discrete time state space model for the system. In digital control systems discrete time models are very important. Discrete time models are also very important for simulation purposes of a dynamic system.

Consider now the case where we let t_0 in (B.16) and (B.17) take the discrete time values

$$t_0 = k\Delta t \quad \forall \quad k = 0, 1, \ldots, \tag{B.18}$$

We then have a discrete time model of the form

$$x((k + 1)\Delta t) = \Phi(\Delta t)x(k\Delta t) + \Delta u(k\Delta t), \tag{B.19}$$

It is however common to use the short hand notation

$$x_{k+1} = \Phi x_k + \Delta u_k. \tag{B.20}$$

We have here defined

$$x_k = x(k\Delta t) = x(t_0) \tag{B.21}$$

Note also that we usually are using symbols as A and B also for discrete time state space models, e.g., so that the model (B.20) is written as

$$x_{k+1} = Ax_k + Bu_k. \tag{B.22}$$

we will usually using the symbols A and B also for discrete time models. However, in cases where there can be conflicts symbols as Φ and Δ are used.

It is important to note that the steady state solution to a continuous time model $\dot{x} = Ax + Bu$ can be found by putting $\dot{x} = 0$. I.e., the steady state solution when time approach infinity ($t \to \infty$) is given by

$$x = -A^{-1}Bu. \tag{B.23}$$

Here the system matrix A is assumed to be non singular.

In a stable system, the transients and dynamic responses will die out as time approach infinity, and al variables will be constant as function of time. Therefore is also the derivative of the states with time equal to zero, i.e.

$$\dot{x} = \frac{dx}{dt} = 0 \tag{B.24}$$

Note also that the steady state solution of a continuous time model and the discrete time model should be the same. This is obvious

B.2.1 Proof of the solution of the state equation

It can be shown that the homogenous solution to the state equation $\dot{x} = Ax+Bu$ (with known initial state $x(t_0)$) when $u = 0$ is of the form

$$x(t) = e^{A(t-t_0)}z \tag{B.25}$$

because $dx/dt = Ae^{A(t-t_0)}z = Ax$.

The solution of the in homogenous differential equation can be found by assuming that the vector z is time variant. We then have from (B.25) that

$$\dot{x} = Ae^{A(t-t_0)}z + e^{A(t-t_0)}\dot{z} \tag{B.26}$$

We also have from the state equation that

$$\dot{x} = Ax + Bu = Ae^{A(t-t_0)}z + Bu \tag{B.27}$$

where we have used that x is given as in (B.25). Comparing (B.26) and (B.27) shows that

$$e^{A(t-t_0)}\dot{z} = Bu. \tag{B.28}$$

This gives

$$\dot{z} = e^{-A(t-t_0)}Bu \tag{B.29}$$

where we have used that $(e^A)^{-1} = e^{-A}$. This gives the following solution for the vector z, i.e.,

$$z(t) = z_0 + \int_{t_0}^{t} e^{-A(\tau-t_0)}Bud\tau. \tag{B.30}$$

We are putting (B.30) in (B.25). This gives

$$x(t) = e^{A(t-t_0)}z = e^{A(t-t_0)}z_0 + \int_{t_0}^{t} e^{A(t-\tau)}Bud\tau. \tag{B.31}$$

Putting $t = t_0$ shows that $x(t_0) = z_0$ and the final solution is found to be

$$x(t) = e^{A(t-t_0)}x(t_0) + \int_{t_0}^{t} e^{A(t-\tau)}Bud\tau. \tag{B.32}$$

This section is meant to show how the solution to the continuous time state equation can be proved.

B.3 Linear transformation of state space models

Let x be the state vector in the state space realization (A, B, D) such that

$$\dot{x} = Ax + Bu, \qquad x(0) = x_0 \tag{B.33}$$
$$y = Dx \tag{B.34}$$

An equivalent realization of the system defined by (B.33) and (B.34) can be found by choosing another basis for the state space (choosing another state). The state vector x can be transformed to a new coordinate system. This can be done by defining a non singular transformation matrix $T \in \mathbb{R}^{n\times n}$ and the following linear transformation,

$$x = Tz \quad \Leftrightarrow \quad z = T^{-1}x \tag{B.35}$$

An equivalent realisation of the system (B.33) and (B.34) is then given by

$$\dot{z} = T^{-1}ATz + T^{-1}Bu, \qquad z(0) = T^{-1}x_0 \tag{B.36}$$
$$y = DTz \tag{B.37}$$

These two state space realizations can be shown to be identical and represent the same system. The two systems has the same transfer function from the input, u to the output, y.

An infinite number of non singular transformation matrices T can be chosen. This leads to an infinite number of state space model realizations. Some of these realizations has special properties, e.g., state space models with special properties can be found by choosing T properly.

B.4 Eigenvalues and eigenvectors

Considder given a matrix $A \in \mathbb{R}^{n\times n}$. The characteristic polynomial of A is then defined as

$$\begin{aligned} p(A) &= \det(\lambda I - A) = \det(A - \lambda I) && \text{(B.38)} \\ &= \lambda^n + p_n\lambda^{n-1} + \cdots + p_2\lambda + p_1 && \text{(B.39)} \end{aligned}$$

where the n coefficients $p_1, p_2, \ldots, p_{n-1}, p_n$ are real values. These coefficients can be found by actually expressing the determinant. The *characteristic equation* is defined as

$$p(A) = \det(\lambda I - A) = 0. \tag{B.40}$$

The n roots of the character polynomial and eqivalently, the n solutions to the characteristic equation, $\lambda_i \ \forall \ i = 1, \ldots, n$, is defined as the eigenvalues of the matrix A. The characteristic equation has always n solutions and the matrix $A \in \mathbb{R}^{n\times n}$ has always n eigenvalues. The eigenvalues can be real or complex. If the eigenvalues are complex, then they will consists of complex conjugate pair, i.e., if $\lambda_k = \alpha + j\beta$ is an eigenvalue, then $\lambda_{k+1} = \alpha - j\beta$ will also be an eigenvalue. The matrix A is said to have multiple eigenvalues if two or more of the eigenvalues are identical.

The spectrum of the eigenvalues of the matrix A is defined as all the eigenvalues, i.e. the collection $\sigma(A) := \{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ of al the eigenvalues is the spectrum of A.

The spectral radius of the matrix A is defined from the eigenvalue with the largest absolute value, i.e., $\rho(A) = \max|\lambda_i| \ \forall \ i = 1, \ldots, n$.

The following theorem is useful in linear algebra and system theory

Theorem B.4.1 (Cayley-Hamilton) *Given a matrix $A \in \mathbb{R}^{n\times n}$ and the characteristic polynomial*

$$p(A) = det(\lambda I - A) = \lambda^n + p_n\lambda^{n-1} + \cdots + p_2\lambda + p_1. \tag{B.41}$$

The Cayley-Hamilton theorem states that the matrix A satisfies its own characteristic polynomial, i.e., such that

$$A^n + p_nA^{n-1} + \cdots + p_2A + p_1I = 0 \tag{B.42}$$

where I is the identity matrix with the same dimension as A.

Proof The prrof is stated only for the case in which the eigenvector matrix M of A is non-singular. An eigenvalue decomposition of A gives that $A = M\Lambda M^{-1}$ where Λ is a diagonal matrix with the eigenvalues $\lambda_i \ \forall \ i = 1, \ldots, n$ on the diagonal. Putting this into (B.42) gives

$$M(\Lambda^n + p_n\Lambda^{n-1} + \cdots + p_2\Lambda + p_1 I)M^{-1} = 0. \tag{B.43}$$

This gives n equations, i.e.,

$$\lambda_i^n + p_n\lambda_i^{n-1} + \cdots + p_2\lambda_i + p_1 = 0 \ \forall \ i = 1, \ldots, n \tag{B.44}$$

which is of the same form as the characteristic equation (B.41). $\triangle$

The Cayley-Hamilton theorem can be used for, e.g.:

- to find an expression for the matrix inverse A^{-1}, i.e.

$$A^{-1} = -\frac{1}{p_1}(A^{n-1} + p_n A^{n-2} + \cdots + p_2 I) \tag{B.45}$$

- to find an expression for the power A^n as a function of $A^{n-1}, \cdots, A$, i.e.

$$A^n = -(p_n A^{n-1} + \cdots + p_2 A + p_1 I) \tag{B.46}$$

- to find a way of computing the coefficients $p_1, \cdots, p_n$ of the characteristic polynomial by Krylovs method. This is presented in the next section.
- develop the controllability and observability matrices of an linear dynamical system, i.e. from the matrix pairs (A, B) and (A, D).

B.4.1 Krylovs method used to find the coefficients of the characteristic equation

We will in this section study a method which can be used to compute the coefficients, $p_1, \cdots, p_n$, in the characteristic polynomial of a $n \times n$ matrix A. This method is referred to as Krylovs method, Krylov (1931).

If we multiply Equation (B.42) from right with a vector $b \in \mathbb{R}^n$, then a linear system of equations can be defined as

$$\overbrace{\begin{bmatrix} b & Ab & \cdots & A^{n-1}b \end{bmatrix}}^{C_n} \overbrace{\begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix}}^{p} = -A^n b. \tag{B.47}$$

This equation, $C_n p = -A^n b$, can be solved with respect to the vector p of coefficients. We have that

$$p = -C_n^{-1} A^n b \tag{B.48}$$

if the vector b is chosen in such a way that the matrix pair (A, b) is controllable, i.e.. in this case in such a way that the controllability matrix C_n is invertible (non-singular).

An arbitrarily random vector b is here usually sufficient. Note that the solution p generally is independent of the choice of b as long as the matrix C_n is invertible.

Note also that (B.42) can directly be written as the linear system of equations

$$\mathcal{A}p = -\text{vec}(A^n) \tag{B.49}$$

where

$$\mathcal{A} = \begin{bmatrix} \text{vec}(I) & \text{vec}(A) & \cdots & \text{vec}(A^{n-1}) \end{bmatrix} \in \mathbb{R}^{n^2 \times n}. \tag{B.50}$$

The solution is given by

$$p = -(\mathcal{A}^T\mathcal{A})^{-1}\mathcal{A}^T\text{vec}(A^n) \tag{B.51}$$

An advantage of this method is that we do not have to chose the vector b. A disadvantage is that this last method is much more computing expense than the first method in which an arbitrarily vector b is chosen.

B.5 Similarity Transformations and eigenvectors

Assume given a non-singular matrix

$T \in \mathbb{R}^{n\times n}$ and a matrix $A \in \mathbb{R}^{n\times n}$. The matrix B defined by

$$B = T^{-1}AT \tag{B.52}$$

is then said to be similar to A. In particular, the eigenvalues of B is identical to the eigenvalues of A. The equation (B.52) is defined as a similarity transformation.

If we are putting the transformation matrix T equal to the eigenvector matrix, M, of the matrix A, then we have that

$$\Lambda = M^{-1}AM \tag{B.53}$$

where Λ is the eigenvalue matrix of the system (matrix A). The eigenvalue matrix Λ os a diagonal matrix with the eigenvalues on the diagonal. This can equivalently be written as

$$AM = M\Lambda \tag{B.54}$$

where

$$M = \begin{bmatrix} m_1 & m_2 & \cdots & m_n \end{bmatrix} \tag{B.55}$$

The columns, m_1, $\cdots$, m_n in the eigenvector matrix M is the eigenvectors corresponding to the eigenvalues λ_1, $\cdots$, λ_n.

Λ is a diagonal matrix with the eigenvalues $\lambda_1, \cdots, \lambda_n$ on the diagonal. Hence, Equation (B.54) can then be written as n linear equations which can be used to compute the eigenvectors, i.e.,

$$\begin{aligned} Am_1 &= \lambda_1 m_1 \\ Am_2 &= \lambda_2 m_2 \\ &\vdots \\ Am_n &= \lambda_n m_n \end{aligned} \tag{B.56}$$

If the matrix A and the eigenvalues $\lambda_1, \cdots, \lambda_n$ are known, then, the eigenvectors and eigenvector matrix (B.55) can be found by solving the linear equations (B.56).

B.6 Time constant

Consider the case in which a 1st order differential equation

$$\dot{x} = ax, \tag{B.57}$$

with known initial value $x(t_0)$ is given. The time constant of the system is then defined as

$$T = -\frac{1}{a}. \tag{B.58}$$

The solution to Equation (B.57) can then be written as

$$x(t) = e^{a(t-t_0)}x(t_0) = e^{-\frac{1}{T}(t-t_0)}x(t_0). \tag{B.59}$$

We see that the solution $x(t)$ at time instant $t = t_0 + T$ is given by

$$x(t_0 + T) = e^{-\frac{1}{T}(t_0+T-t_0)}x(t_0) = e^{-1}x(t_0) \approx 0.37x(t_0). \tag{B.60}$$

I.e., the sesponse have fallen 63%.

Definition B.1 (Time constants)
Given a linear dynamic system, $\dot{x} = Ax + Bu$, where $x \in \mathbb{R}^n$ is the state vector of the system. The system matrix A has n eigenvalues given by

$$\lambda_i = \lambda_i(A) \ \forall \ i = 1, \ldots, n. \tag{B.61}$$

If the eigenvalues are all real, distinct and have negative values (stable system), then the system will have th n time constants given by

$$T_i = -\frac{1}{\lambda_i} \ i = 1, \ldots, n. \tag{B.62}$$

Note also that the connection with the eigenvalues in a discrete time system

$$x_{k+1} = \phi x + \delta u, \tag{B.63}$$

and the continuous equivalent

$$\dot{x} = -\frac{1}{T}x + bu, \tag{B.64}$$

then is given by

$$\phi = e^{-\frac{1}{T}\Delta t} \tag{B.65}$$

which gives that

$$T = -\frac{\Delta t}{\ln \phi}. \tag{B.66}$$

Methods in system identification can be used to identify discrete time models from known input and output data of a system. Usually there are the parameters ϕ and δ which are estimated (computed). The relationship (B.66) is therefore very useful in order to find the time constant of the real time system.

B.7 The matrix exponent and the transition matrix

We have earlier in this section shown that that the transition matrix are involved in the exact solution of a linear time invariant dynamical system. Considder the autonomous system

$$\dot{x} = Ax \tag{B.67}$$

with known initial value $x_0 = x(t_0 = 0)$. Then the solution is given by

$$x(t) = \Phi(t)x(0) \tag{B.68}$$

where the transition matrix $\Phi(t)$ is given by

$$\Phi(t) = e^{At}. \tag{B.69}$$

As we see, the problem of computing the transition matrix $\Phi(t)$, is the same problem as computing the matrix exponent

$$F = e^{A}. \tag{B.70}$$

B.7.1 Computing the matrix exponent by diagonalisation

Let $f(A)$ be an analytical matrix function of A which also should contain the eigenvalue spectrum of A. A more general formulation of the similarity transformation given in (B.52) is then defined as

$$f(B) = T^{-1}f(A)T \tag{B.71}$$

Assume now that we want to compute the matrix exponent e^A. As we have shown in Equation (B.53 the matrix A can be decomposed as

$$A = M\Lambda M^{-1} \tag{B.72}$$

when the eigenvector matrix M is invertible. Using (B.72), (B.71) and $f(A) = e^A$ gives

$$e^A = Me^{\Lambda}M^{-1} \tag{B.73}$$

As we see, when the eigenvector matrix M and the eigenvalue matrix Λ of the matrix A are known, then the matrix exponential e^A can simply be computed from (B.73).

Equation (B.73) can be proved by starting with the autonomous system

$$\dot{x} = Ax \tag{B.74}$$

with known initial state vector $x(0)$. This system has the solution

$$x(t) = e^{At}x(0). \tag{B.75}$$

Transforming (B.74) by using $x = Mz$ gives

$$\dot{z} = \Lambda z \tag{B.76}$$

with initial state $z(0) = M^{-1}x(0)$. The canonical (transformed) system (B.76) have the solution

$$z(t) = e^{\Lambda t}z(0). \tag{B.77}$$

Using the transformation $x = Mz$ and putting this into (B.77) gives

$$x(t) = Me^{\Lambda t}z(0) = Me^{\Lambda t}M^{-1}x(0). \tag{B.78}$$

Comparing the two solutions (B.78) and (B.75) gives (B.73).

Note that in some circumstances there may be simpler to compute the transition matrix or matrix exponential $f(A)$ by solving the equation system

$$f(A)M = Mf(\Lambda) \tag{B.79}$$

because we in this case do not explicitly have to compute the matrix inverse M^{-1}. For some problems M is not invertible. This may be the case for systems which have multiple eigenvalues. We are referring to Parlet (1976) for a more detailed description of matrix functions and computing methods.

B.7.2 Parlets method for computing the matrix exponent

It can be shown, Parlet (1976), that the matrix exponent $F = e^A$ and the system matrix A commutes, i.e. the following is satisfied

$$FA = AF \tag{B.80}$$

If the matrix A has a special structure, e.g., upper or lower triangular, then Equation (B.7.2) can with advantage be used in order to compute the unknown elements in the transition matrix.

Note that the matrix exponential of an upper triangular matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix}, \tag{B.81}$$

is given by

$$F = e^A = \begin{bmatrix} e^{a_{11}} & f_{12} \\ 0 & e^{a_{22}} \end{bmatrix}. \tag{B.82}$$

The unknown coefficient f_{12} can then simply be found from equation ().

Example B.2 (computing the transition matrix)
Given an autonomous system described by

$$\dot{x} = Ax, \tag{B.83}$$

with the initial state $x_0 = x(0)$. The system matrix A is given by

$$A = \begin{bmatrix} \lambda_1 & \alpha \\ 0 & \lambda_2 \end{bmatrix}. \tag{B.84}$$

We want to compute the transition matrix

$$F = e^{At} \tag{B.85}$$

by Parlets method. First we find immediately that

$$F = \begin{bmatrix} e^{\lambda_1 t} & f \\ 0 & e^{\lambda_2 t} \end{bmatrix}. \tag{B.86}$$

We now have to find the unknown constant f_{12} in the transition matrix. This can be done from the equation system

$$AF = FA. \tag{B.87}$$

This gives four equations but only one of them gives information of interest,

$$\lambda_1 f_{12} + \alpha e^{\lambda_2 t} = e^{\lambda_1 t}\alpha + f_{12}\lambda_2. \tag{B.88}$$

Solving with respect to f_{12} gives

$$f_{12} = \alpha \frac{e^{\lambda_1 t} - e^{\lambda_2 t}}{\lambda_1 - \lambda_2} \tag{B.89}$$

As we see, this method can simply be used for system matrices which have a triangular structure, and in which the eigenvalues are distinct and not identical to zero.

B.7.3 Matrix exponential by series expansion

It can be shown that the matrix exponential F^A can be expressed as an infinite Taylor series

$$e^A = I + A + \frac{1}{2}A^2 + \cdots \tag{B.90}$$

The transition matrix can be expressed in the same way, e.g.,

$$e^{At} = I + At + \frac{1}{2}A^2t^2 + \cdots \tag{B.91}$$

This is in general not a good method for computing the transition matrix, because it will in general lead to numerical problems when computing powers of A like A^9, A^{10}, etc. especially when A contains small values. This is so due to the finite precision of the computer. Note that the machine precision of a 32 bit computer is $eps = 1/2^{52}$.

The series method is however very useful for computing the transition matrix of many simple systems. This will be illustrated in the following example.

Example B.3 (computing transition matrix)
Given an autonomous system described by

$$\dot{x} = Ax, \tag{B.92}$$

where the initial state $x_0 = x(0)$ is given and the system matrix is given by

$$A = \begin{bmatrix} 0 & \alpha \\ 0 & 0 \end{bmatrix}. \tag{B.93}$$

The transition matrix for this system is simply found from the two first terms of the Taylor series (B.91) because A is so called nil-potent, i.e., we have that $A^2 = 0$, $A^3 = 0$ and so on. We have

$$\Phi(t) = I + At = \begin{bmatrix} 1 & \alpha t \\ 0 & 1 \end{bmatrix}. \tag{B.94}$$

B.8 Examples

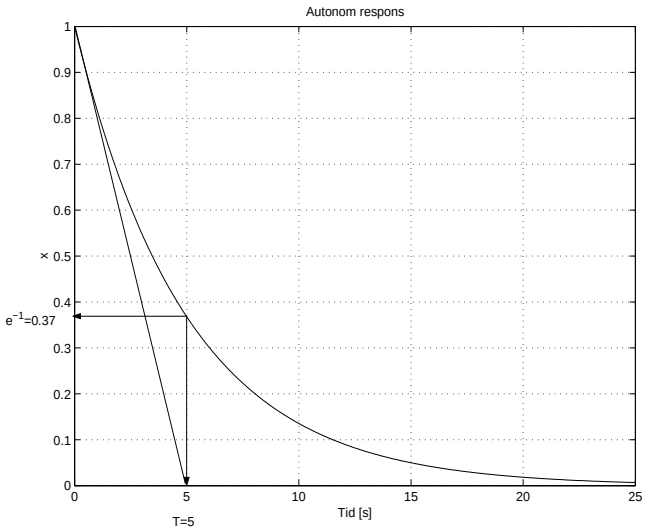
Example B.4 (autonomous response and time constant)
Given an autonomous system

$$\dot{x} = ax \tag{B.95}$$

where the initial state is $x_0 = x(t_0 = 0) = 1$ and the system parameter $a = -\frac{1}{T}$ where the time constant is $T = 5$. The solution of this differential equation is

$$x(t) = e^{-\frac{1}{T}t}x_0 = e^{-\frac{1}{5}t}. \tag{B.96}$$

Let us now plot the solution in the time interval $0 \leq t \leq 25$. Note that the state will have approximately reached the steady state value after $4T$ (four times the time constant). The solution is illustrated in Figure B.1.



Figur B.1: Tidsrespons av det autonome systemet $\dot{x} = ax$ der $x_0 = 0$ og $a = -\frac{1}{T}$ for $T = 5$.

Example B.5 (computation of matrix exponent)
Given the system matrix

$$A = \begin{bmatrix} -3 & 1 \\ 2 & -2 \end{bmatrix}. \tag{B.97}$$

The eigenvalue matrix Λ *and the corresponding eigenvector matrix can be shown to be as follows*

$$M = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} -1 & 0 \\ 0 & -4 \end{bmatrix} \tag{B.98}$$

Find the matrix exponent $F = e^A$.

We have the relation $F = Me^{\Lambda}M^{-1}$, *which is equivalent with*

$$FM = Me^{\Lambda} \tag{B.99}$$

which gives

$$\begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} e^{-1} & 0 \\ 0 & e^{-4} \end{bmatrix} \tag{B.100}$$

From this we have the four equations

$$\begin{bmatrix} f_{11} + 2f_{12} & f_{11} - f_{12} \\ f_{21} + 2f_{22} & f_{21} - f_{22} \end{bmatrix} = \begin{bmatrix} e^{-1} & e^{-4} \\ 2e^{-1} & -e^{-4} \end{bmatrix} \tag{B.101}$$

Taking element $1,1$ *minus element* $1,2$ *on the left hand side of Equation (B.101) gives*

$$3f_{12} = e^{-1} - e^{-4} \tag{B.102}$$

Putting the expression for f_{12} into element $1,2$ on the left hand side gives f_{11}, i.e.,

$$f_{11} - f_{12} = e^{-4} \tag{B.103}$$

which gives

$$f_{11} = \frac{1}{3}(e^{-1} + 2e^{-4}) \tag{B.104}$$

Taking element $2,1$ minus element $2,2$ on the left hand side of (B.101) gives

$$3f_{22} = 2e^{-1} + e^{-4} \tag{B.105}$$

Putting the expression for f_{22} into e.g., element $2,2$ gives f_{21}. This gives the final result

$$F = e^{A} = \frac{1}{3}\begin{bmatrix} e^{-1} + 2e^{-4} & e^{-1} - e^{-4} \\ 2e^{-1} - 2e^{-4} & 2e^{-1} + e^{-4} \end{bmatrix} \tag{B.106}$$

Note that the transition matrix could have been computed similarly, i.e.,

$$\Phi(t) = e^{At} = Me^{\Lambda t}M^{-1} = \frac{1}{3}\begin{bmatrix} e^{-t} + 2e^{-4t} & e^{-t} - e^{-4t} \\ 2e^{-t} - 2e^{-4t} & 2e^{-t} + e^{-4t} \end{bmatrix} \tag{B.107}$$

Example B.6 (computation of transition matrix for upper triangular system)

Consider given an autonomous system described by the matrix differential equation

$$\dot{x} = Ax, \tag{B.108}$$

where the initial state $x_0 = x(0)$ is given and where the system matrix is given by

$$A = \begin{bmatrix} \lambda_1 & \alpha \\ 0 & \lambda_2 \end{bmatrix}. \tag{B.109}$$

The transition matrix $\Phi(t) = e^{At}$ will have the same upper triangular strukture as A and the diagonal elements in $\Phi(t)$ is simply $e^{-\lambda_1 t}$ and $e^{-\lambda_2}$, i.e.,

$$\Phi(t) = \begin{bmatrix} e^{\lambda_1 t} & f_{12} \\ 0 & e^{\lambda_2 t} \end{bmatrix}. \tag{B.110}$$

The unknown element f_{12} can now simply be computed from Parlets method, i.e., we solve the equation

$$\Phi At = At\Phi \tag{B.111}$$

or equivalent

$$\Phi A = A\Phi. \tag{B.112}$$

This gives the equation

$$e^{\lambda_1 t}\alpha + f_{12}\lambda_2 = \lambda_1 f_{12} + \alpha e^{\lambda_2 t}. \tag{B.113}$$

Solving for the remaining element f_{12} gives

$$f_{12} = \alpha \frac{e^{\lambda_1 t} - e^{\lambda_2 t}}{\lambda_1 - \lambda_2}. \tag{B.114}$$

Note that this method only can be used when the system has distinct eigenvalues, i.e. when $\lambda_1 \neq \lambda_2$.

However, it can be shown that in the limit when $\lambda_2 \to \lambda_1 = \lambda$ that

$$f_{12} = \alpha t e^{\lambda_1 t} \tag{B.115}$$

Example B.7 (Set of higher order ODE to set of first order ODE)
Consider a system described be the following couple of coupled differential equations

$$\ddot{y}_1 + k_1\dot{y}_1 + k_2 y_1 = u_1 + k_3 u_2$$
$$\dot{y}_2 + k_4 y_2 + k_3\dot{y}_1 = k_6 u_1$$

where u_1 and u_2 is defined as the control inputs and y_1 and y_2 is defined as the measurements or outputs

We now define the outputs and if necessary the derivatives of the outputs as states. Hence, define the states

$$x_1 = y_1, \quad x_2 = \dot{y}_1, \quad x_3 = y_2 \tag{B.116}$$

This gives the following set of 1st order differential equations for the states

$$\dot{x}_1 = x_2 \tag{B.117}$$
$$\dot{x}_2 = -k_2 x - 1 - k_1 x_2 + u_1 + k_3 u_2 \tag{B.118}$$
$$\dot{x}_3 = -k_5 x_2 - k_4 x_3 + k_6 u_1 \tag{B.119}$$

and the following measurements (outputs) variables

$$y_1 = x_1 \tag{B.120}$$
$$y_2 = x_3 \tag{B.121}$$

The model is put on matrix (State Space) form as follows

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -k_2 & -k_1 & 0 \\ 0 & -k_5 & k_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & k_3 \\ k_6 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_3 \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

and finally in matrix form as follows

$$\dot{x} = Ax + Bu \tag{B.122}$$
$$y = Dx \tag{B.123}$$

B.9 Transfer function and transfer matrix

Laplace transforming $\dot{x}$ gives

$$L(\dot{x}(t)) = sx(s) - x(t=0). \tag{B.124}$$

Similarly, the Laplace transform of a time dependent variable x is defined as

$$L(x(t)) = x(s). \tag{B.125}$$

Using (B.124) and the definition (B.125) in the state space model

$$\begin{aligned} \dot{x} &= Ax + Bu, \qquad x(t=0) = 0, && \text{(B.126)} \\ y &= Dx + Eu, && \text{(B.127)} \end{aligned}$$

gives the Laplace transformed model equivalent

$$\begin{aligned} x(s) &= (sI - A)^{-1}Bu(s), && \text{(B.128)} \\ y(s) &= Dx(s) + Eu(s). && \text{(B.129)} \end{aligned}$$

We can now write (B.128) and (B.129) as a transfer matrix model

$$y(s) = H(s)u(s), \tag{B.130}$$

where $H(s)$ is the transfer matrix of the system

$$H(s) = D(sI - A)^{-1}B + E. \tag{B.131}$$

For single-input and single-output systems then $H(s)$ will be a scalar function of the Laplace variable s. In this case we usually are using a small letter, i.e., we are putting $h(s) = H(s)$. Note also that we have included a direct influence from the input u to the output y in the measurement (output) equation- This will be the case in some circumstances. However, the matrix or parameter E is usually zero in control systems.

Note also that when the eigenvalue decomposition $A = M\Lambda M^{-1}$ exists then we have that the transfer matrix can be expressed and computed as follows

$$\begin{aligned} H(s) &= D(sI - A)^{-1}B + E && \text{(B.132)} \\ &= DM(sI - \Lambda)^{-1}M^{-1}B + E. && \text{(B.133)} \end{aligned}$$

Finally note the following important relationship. The properties of a time dependent state space model when $t \to \infty$, i.e. the steady state properties, can be analysed in a transfer function Laplacian model by putting $s = 0$. The transient behaviour when $t = 0$ is analysed by letting $s \to \infty$.

B.10 Linearization

In many cases the starting point of a control problem or model analysis problem is a non-linear model of the form

$$\dot{x} = f(x,u), \tag{B.134}$$
$$y = g(x,u). \tag{B.135}$$

Here $x \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}^r$ is the control input vector and $y \in \mathbb{R}^m$ is the output or measurements vector. The functions $f(\cdot,\cdot) \in \mathbb{R}^n$ and $g(\cdot,\cdot) \in \mathbb{R}^m$ may be non-linear smooth functions of x and u. Note also that the initial state is $x(t_0)$ which should be given ore known before the state space model can be simulated in time.

In this case it may be of interest to derive a linear model approximation to (B.134) and (B.135).

The two first (linear terms) of a Taylor series expansion of the right hand side of (B.134) around the points x_0 and u_0 gives

$$f(x,u) \approx f(x_0,u_0) + \left.\frac{\partial f}{\partial x^T}\right|_{x_0,u_0}(x-x_0) + \left.\frac{\partial f}{\partial u^T}\right|_{x_0,u_0}(u-u_0). \tag{B.136}$$

Define the deviation variables

$$\Delta x = x - x_0, \tag{B.137}$$
$$\Delta u = u - u_0. \tag{B.138}$$

Also define the matrices

$$A = \left.\frac{\partial f}{\partial x^T}\right|_{x_0,u_0} \in \mathbb{R}^{n\times n} \tag{B.139}$$

which also is named the Jacobian matrix. Similarly, define

$$B = \left.\frac{\partial f}{\partial u^T}\right|_{x_0,u_0} \in \mathbb{R}^{n\times r} \tag{B.140}$$

Putting (B.136), (B.137) and (B.138) into the state equation (B.134) gives the linearized state equation model

$$\dot{\Delta x} = A\Delta x + B\Delta u + v, \tag{B.141}$$

where

$$v = f(x_0,u_0) - \dot{x}_0. \tag{B.142}$$

Usually the points x_0 and u_0 is constant steady state values such that

$$\dot{x_0} = f(x_0,u_0) = 0. \tag{B.143}$$

Hence, a linearized state equation is given by

$$\dot{\Delta x} = A\Delta x + B\Delta u. \tag{B.144}$$

Similarly the output equation (B.135) can be linearized by approximating the right hand side by the first two terms of a Taylor series expansion, i.e.,

$$y \approx g(x_0, u_0) + \left.\frac{\partial g}{\partial x^T}\right|_{x_0,u_0} (x - x_0) + \left.\frac{\partial g}{\partial u^T}\right|_{x_0,u_0} (u - u_0). \tag{B.145}$$

Now defining

$$\begin{aligned} y_0 &= g(x_0, u_0) && \text{(B.146)}\\ \Delta y &= y - y_0 && \text{(B.147)}\\ D &= \left.\frac{\partial g}{\partial x^T}\right|_{x_0,u_0} && \text{(B.148)}\\ E &= \left.\frac{\partial g}{\partial u^T}\right|_{x_0,u_0} && \text{(B.149)} \end{aligned}$$

gives the linearized output equation

$$\Delta y = D\Delta x + E\Delta u. \tag{B.150}$$

Usually the deviation variables are defined as

$$\begin{aligned} x &:= x - x_0, && \text{(B.151)}\\ u &:= u - u_0. && \text{(B.152)} \end{aligned}$$

Hence, a linear or linearized state space model, given by (B.144) and (B.150), is usually written as follows.

$$\begin{aligned} \dot{x} &= Ax + Bu, && \text{(B.153)}\\ y &= Dx + Eu. && \text{(B.154)} \end{aligned}$$

One should therefore note that the variables in a linearized model may be deviation variables, but this is not always the case. One should also note that only linear models can be transformed to Laplace plane models. Note also that the initial state in the linearized model is given by $\Delta x(t_0) = x(t_0) - x_0$.

Example B.8 (Linearization of a pendulum model)
An non-linear model for a pendulum can be written as the following second order differential equation, i.e.,

$$\ddot{\theta} + + \frac{b}{mr^2}\dot{\theta} + \frac{g}{r}\sin(\theta) = 0, \tag{B.155}$$

where θ is the angular position (deviation of the pendulum from the vertical line, i.e. from the steady state position). $m = 8$ is the mass of the pendulum, $r = 5$ is the length of the pendulum arm, $b = 10$ is a friction coefficient in the base point and $g = 9.81m/s^2$ is the acceleration of gravity constant.

The second order model can be written as a set of 1st order differential equations by defining the states

$$\begin{aligned} x_1 &= \theta \quad \textit{angular position} && \text{(B.156)}\\ x_2 &= \dot{\theta} \quad \textit{angular velocity} && \text{(B.157)} \end{aligned}$$

from this definitions we have that $\dot{x}_1 = \dot{\theta} = x_2$ *which gives the state space model*

$$\dot{x}_1 = x_2, \tag{B.158}$$

$$\dot{x}_2 = -\frac{g}{r}\sin(x_1) - \frac{b}{mr^2}x_2, \tag{B.159}$$

which is equivalent to a non-linear model

$$\dot{x} = f(x) \tag{B.160}$$

with the vector function

$$f(x) = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -\frac{g}{r}\sin(x_1) - \frac{b}{mr^2}x_2 \end{bmatrix} \tag{B.161}$$

Linearizing around the steady state solution $x_1 = 0$ *and* $x_2 = 0$ *gives*

$$\dot{x} = Ax, \tag{B.162}$$

where the Jacobian is given by

$$A = \left.\frac{\partial f}{\partial x^T}\right|_0 = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix}_0 = \begin{bmatrix} 0 & 1 \\ -\frac{g}{r}\cos(x_1) & -\frac{b}{mr^2} \end{bmatrix}_0 = \begin{bmatrix} 0 & 1 \\ -\frac{g}{r} & -\frac{b}{mr^2} \end{bmatrix} \tag{B.163}$$

Putting into the numerical values we obtain

$$A = \begin{bmatrix} 0 & 1 \\ -1.962 & -0.050 \end{bmatrix} \tag{B.164}$$

Note that the linearized model could have been obtained more directly by using that $\sin(x_1) \approx x_1$ *for small angles* x_1*.*

Example B.9 (Simulation of a non-linear pendulum model)
The nonlinear state space pendulum model

$$\dot{x}_1 = x_2, \tag{B.165}$$

$$\dot{x}_2 = -\frac{g}{r}\sin(x_1) - \frac{b}{mr^2}x_2, \tag{B.166}$$

with $g = 9.81$*,* $r = 5$*,* $m = 8$ *and* $b = 10$ *can be simply simulated in MATLAB by using an ODE solver, e.g.,*

```
>> sol=ode15s(@ fx_pendel, 0:0.1:50,[1;0]);
>> plot(sol.x,sol.y)
```

Here sol is an object where sol.x is the time axis and sol.y is the states. Try it! The file fx_pendel is an m-file function given in the following

```
function fx=fx_pendel(t,x)
% fx_pendel
% fx=fx_pendel(t,x)
% Modell av pendel.

m=8; g=9.81; b=10; r=5;

fx=zeros(2,1);
fx(1)=x(2);
fx(2)=-b*x(2)/(m*r^2)-g*sin(x(1))/r;
```