
Linjär Algebra V

2015



Fraunhofer

CHALMERS

Research Centre

Industrial Mathematics

Johan Karlsson

`johan.karlsson@fcc.chalmers.se`

1.1 Linjära ekvationssystem

Linjär ekvation: $\sum_{i=1}^n a_i x_i = b$, a_i *koefficienter (reella tal)*

Flera ekvationer kan forma ett system:

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

m ekvationer, n obekanta x

Ofta bra att samla a_{ij} och b_i i matriser.

1.1 Matriser

Rektangulär uppsättning reella tal (skalärer).

En $m \times n$ -matris har m rader och n kolumner

Matrisaddition, subtraktion & multiplikation med skalär elementvis.

Matrismultiplikation (kommuterar ej, $AB \neq BA$)

$$A \cdot B = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1k} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nk} \end{bmatrix} =$$

$$\begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix} [B_1 \quad \cdots \quad B_k] = \begin{bmatrix} A_1 B_1 & \cdots & A_1 B_k \\ \vdots & \ddots & \vdots \\ A_m B_1 & \cdots & A_m B_k \end{bmatrix}$$

$$A_i B_j = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{in} b_{nj}$$

Notera: n matchar i A och B & AB är $m \times k$.

Notera: $A[B_1 \quad \cdots \quad B_k] = [AB_1 \quad \cdots \quad AB_k]$

En matris med 1or på diagonalen och nollor i övrigt kallas identitetsmatris (enhetsmatris) och betecknas I .



1.1 Linjära ekvationssystem på matrisform

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$
$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = A$$
$$\begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = b$$
$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x$$
$$Ax = b$$



1.1 Lösningar till linjära ekvationssystem

Vi ska lösa linjära ekvations-system $Ax = b$

Ett linjärt ekvations-system kan ha

- En unik lösning x
- Ingen lösning
- Oändligt många lösningar (beskrivs med hjälp av parametrar och kallas parameterlösning)

Två ekvations-system är ekvivalenta om de har samma lösningsmängd.

1.2 Eliminationsmetoden (Gausselimination)

Sätt upp utökade matrisen $[A \ b]$

Omforma till ekvivalent (samma lösningar) system som är enkelt att lösa.

Standardmetod i två steg:

Steg 1: Omforma till övre-triangulärt trappstegs-system med hjälp av elementära radoperationer:

- Skalning av rad med konstant skild från 0.
- Radbyte
- Ersätt rad med raden + multipel av annan rad.

En kolumn i taget från vänster.

Byt rader för att få så triangulärt som möjligt.

Använd en rad för att få nollor i kolumnen i raderna under.

Steg 2: Skapa nästan identitetsmatris till vänster (unik reducerad trappstegsform).

En kolumn i taget från höger.

Normera rad för att få en etta längst till vänster.

Använd raden för att få nollor i kolumnen i raderna över.

Tre fall:

- Om motsägelse så ingen lösning
- Om kolonner utan pivotelement så parameterlösningar
- Annars unik lösning

Räkna nerifrån och upp till lösning(ar).



1.3-1.6 Vektorer

En vektor är en n -tuppel av tal $\in R$ (eller $\in C$).

Mängden av alla vektorer med n element betecknas R^n .

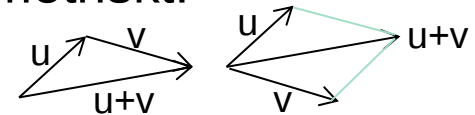
Geometriskt är en vektor en pil med längd och riktning, men utan fast position.

Kan tex representeras med pil från origo till en given punkt.

Vektoraddition beräknas koordinatvis:

$$x * (a, b) + y * (c, d) = (x * a + y * c, x * b + y * d)$$

Geometriskt:



En linjärkombination av vektorerna i $\{v_1, \dots, v_p\}$ är en summa $c_1 v_1 + \dots + c_p v_p$ med skalära vikter $\{c_1, \dots, c_p\}$.

Mängden av alla linjärkombinationer av vektorerna i en mängd $\{v_1, \dots, v_p\}$ (den del av R^n som mängden spänner upp) kallas spannet (linjära höljet) av vektorerna, $\text{span}\{v_1, \dots, v_p\}$



1.3-1.6 $Ax=b$ igen

Ax är linjärkombination av kolumner i A med vikter i x .

EKVIVALENTA:

$Ax=b$ har lösning för varje b i R^m

Varje b i R^m är en linjärkombination av kolumner i A

Kolumnerna i A spänner upp R^m

Radreducerade A har pivotelement i varje rad

1.7 Linjärt beroende

En mängd vektorer $\{v_1, \dots, v_p\}$ i R^n är *linjärt beroende* om det finns skalära vikter $\{c_1, \dots, c_p\}$ (som inte alla är 0) så att

$$c_1 v_1 + \dots + c_p v_p = 0$$

Med andra ord: någon vektor går att uttrycka som en linjärkombination av resten.

Linjärt beroende $\rightarrow$ linjärkombination: lös ut v_i (med $c_i \neq 0$)

Exempel: $1v_1 - 2v_2 + 3v_3 = 0 \Leftrightarrow v_2 = \frac{1}{2}v_1 + \frac{3}{2}v_3$

Linjärkombination $\rightarrow$ linjärt beroende:

flytta över allt till vänster eller högerled (med resten av $c_i = 0$)

Exempel: $v_2 = 3v_1 - v_4 \Leftrightarrow 3v_1 - v_2 + 0v_3 - v_4 = 0$

En mängd vektorer $\{v_1, \dots, v_p\}$ är *linjärt oberoende* om enda lösningen till $c_1 v_1 + \dots + c_p v_p = 0$ är att låta alla c vara 0.



1.7 Linjärt beroende

Om $p > n$ är mängden alltid linjärt beroende (Theorem 8).

$$A = [v_1 \quad \cdots \quad v_p] \quad x = \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix} \quad Ax = 0 \Leftrightarrow c_1 v_1 + \cdots + c_p v_p = 0$$

Homogent system med fler variabler än ekvationer har alltid parameterlösning, dvs finns lösning $x \neq 0$.

Om mängden innehåller nollvektorn är mängden alltid linjärt beroende (Theorem 9).

Om v_i är nollvektorn så sätt $c_i = 1$ och resten av $c_k = 0$

$$c_1 v_1 + \cdots + c_p v_p = 1 v_i = 0$$

1.8 Transformation/avbildning/funktion/mapping

En transformation T från R^n till R^m är en regel som till varje vektor $x \in R^n$ tilldelar en vektor $T(x) \in R^m$

R^n är definitionsmängden (domain)

R^m är målmängden (codomain)

Mängden av alla $T(x)$ kallas värdemängden (range).

En transformation T är linjär omm

$$T(cu + dv) = cT(u) + dT(v)$$

Ett nödvändigt, men inte tillräckligt, krav är $T(0) = 0$

1.9 Matrisen för en transformation

Varje *linjär* transformation $T(x)$ från R^n till R^m är en matris-transformation $T(x) = Ax$ med unikt A .

A har som kolumner avbildningarna (transformationerna) av basvektorerna till R^n . Dessa basvektorer $\{e_1, \dots, e_n\}$ är också kolumnerna i identitetsmatrisen.

$$A = [T(e_1) \dots T(e_n)]$$

$$x = Ix = [e_1 \quad \dots \quad e_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 e_1 + \dots + x_n e_n$$

$$T(x) = x_1 T(e_1) + \dots + x_n T(e_n) = [T(e_1) \quad \dots \quad T(e_n)] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = Ax$$

1.9 One-to-one & Onto

$T: R^n \rightarrow R^m$ är surjektiv (onto) om varje y i R^m
är avbildningen av minst ett x i R^n

$T: R^n \rightarrow R^m$ är injektiv (one to one) om varje y i R^m
är avbildningen av högst ett x i R^n

Sats: T är surjektiv (onto) om A s kolumner spänner upp R^m

Att A s kolumner spänner upp R^m betyder att alla $b \in R^m$ kan skrivas som linjärkombinationer av kolumner i A .

Vilket är samma sak som att $Ax=b$ har lösning för alla b .

Vilket är ekvivalent med definitionen av surjektiv.

1.9 One-to-one & Onto

Sats: T är injektiv (one-to-one) omm $T(x)=0$ bara har den triviala lösningen $x=0$.

$\Rightarrow$

One-to-one betyder att $T(x)=0$ har högst en lösning, och $x=0$ är alltid en lösning till linjära homogena system, alltså inga andra lösningar än $x=0$.

$\Leftarrow$ (motsägelsebevis)

Vi antar alltså att det bara finns lösningen $x=0$.

Antag nu också att T inte är injektiv.

Då finns något y så att både $T(x_1) = y$ och $T(x_2) = y, (x_1 \neq x_2)$.

Då är $T(x_1 - x_2) = T(x_1) - T(x_2) = y - y = 0$ och $x_1 - x_2 \neq 0$.

Dvs det finns inte bara den triviala lösningen, vilken är en motsägelse, alltså måste T vara injektiv.

Sats: T är injektiv (one-to-one) omm A s kolumner är linjärt oberoende.

T inj. omm $T(x)=Ax=0$ bara har den triviala lösningen enligt satsen ovan.

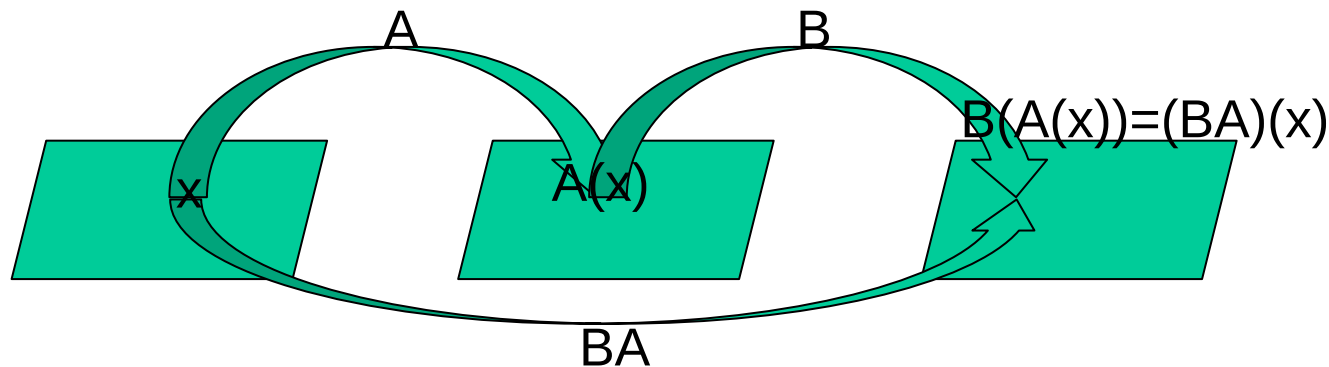
Att $Ax=0$ bara har den triviala lösningen är ekvivalent med definitionen av linjärt oberoende för A s kolumner.



2.1(1) Matrisräkning

En diagonalmatris är en matris med nollor utanför diagonalen $\{a_{ii}\}$

Matrismultiplikation är associativ, vilket betyder att man kan sätta paranteser och utföra multiplikation i vilken ordning man vill, tex $ABC = (AB)C = A(BC)$



2.1(2) Matrisräkning

$$AB \neq BA$$

Matrismultiplikation:

$$AB_{ij} = [A_{i1} \ A_{i2} \ \dots] \begin{bmatrix} B_{1j} \\ B_{2j} \\ \vdots \end{bmatrix} = A_{i1}B_{1j} + A_{i2}B_{2j} + \dots$$

$$A^k = A * \dots * A \text{ (k st), alltså inte elementvis } (a_{ij})^k$$

$$A^0 = I \quad (def)$$

Transponat: rader blir kolonner och tvärtom

$$(A + B)^T = A^T + B^T$$

$$(AB)^T = B^T A^T$$

2.2 Matris-invers

Räkna med matriser som med vanliga tal ungefär,

Men kom ihår att multiplikation inte kommuterar.

Och "division" är speciellt och kräver matrisinvers, vilket inte alltid finns.

Vi skriver alltså inte $\frac{A}{B}$ utan $B^{-1}A$ eller AB^{-1} och de två är inte samma (mult. kommuterar inte).

Tänk på att många operationer egentligen är "division", tex $AB = AC \Leftrightarrow B = C$, vilket alltså bara gäller om invers till A finns,

$AB = 0 \Rightarrow A = 0$ eller $B = 0$, vilket bara gäller om invers till B respektive A finns.

2.2 Matris-invers

A (kvadratisk) är inverterbar om det finns en matris A^{-1} sådan att

$$AA^{-1} = I \quad \text{och} \quad A^{-1}A = I$$

Annars kallas A singular.

Sats: inversen är unik om den finns.

Antag att B & C båda är inverser till A . $B=BI=BAC=IC=C$

Finns invers till A så har $Ax = b$ unik lösning $x = A^{-1}b$ för alla b

$$\text{jfr } ax = b \Leftrightarrow \frac{a}{a}x = \frac{b}{a} \Leftrightarrow x = \frac{b}{a} \quad (\text{om } a \neq 0)$$

$$Ax = b \Leftrightarrow A^{-1}Ax = A^{-1}b \Leftrightarrow x = A^{-1}b \quad (\text{om invers finns})$$

$$(A^T)^{-1} = (A^{-1})^T$$

$$(AB)^{-1} = B^{-1}A^{-1}$$

2.2 Matris-invers

$[A \ I] \sim [I \ A^{-1}]$ Går det inte så är A inte inverterbar.

Elementär radoperation på A kan representeras EA.

E fås genom att göra samma operation på I. Tex

$$E_{\text{skalning rad 3 med } k} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & k \end{bmatrix} \quad E_{\text{radbyte rad 2 \& 3}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad E_{\text{add. } m \cdot \text{rad 3 till rad2}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & m \\ 0 & 0 & 1 \end{bmatrix}$$

Eftersom operation reverserbar så finns invers till E.

Antag A inv.bar. Då är $Ax=b$ lösbar för alla b. Då finns pivotposition i varje rad, och den reducerade formen av A är då I.

Antag $A \sim I$. Då är alltså $E_p \cdots E_1 A = I \Rightarrow A = (E_p \cdots E_1)^{-1}$ dvs A är inv.bar och $A^{-1} = ((E_p \cdots E_1)^{-1})^{-1} = E_p \cdots E_1 = E_p \cdots E_1 I$

För 2x2-matris: $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

Är $ad-bc=0$ så är A inte inverterbar.

OBS: Vill man bara lösa $Ax=b$ så beräkna inte invers utan bara kör radrecucering på utökade matrisen.



2.3-2.4

Följande är ekvivalenta för en $n \times n$ -matris A

- A är inverterbar
- A har n pivotelement
- $Ax=0$ har bara den triviala lösningen $x=0$
- Kolumnerna i A är linjärt oberoende
- $Ax=b$ har alltid lösning

Blockmatriser

- Dela in som man vill i rektangulära block
- Räkna som vanligt med block som element
- En blockdiagonal matris är inverterbar om varje block på diagonalen är inverterbart.

2.5 LU-faktorisering

$A=LU$ $m \times n$

L $m \times m$ vänster/under-triangelär med 1or på diagonalen

U $m \times n$ övretriangelär trappstegsmatris

L beskriver radoperationerna för att göra om A till U (eller egentligen U till A).

Om radbyten krävs för att få A till U så blir inte L triangelär, men "triangelär med permuterade rader".

Varför?

$Ax=b$ $LUx=b$ $Ux=y$ $Ly=b$

Triangelära system snabba att lösa.

Lös först $Ly=b$ för att få y och sedan $Ux=y$ för att få x .

Hur (i fallet utan radbyten)?

Reducera A till U med radoperationer.

L avläses kolumnvis från pivotelement och nedåt innan nollor skapas i den kolumnen under pivotelementet, och sedan normerar vi kolumnvis så att diagonalen får ettor.

2.8 Vektor-rum

Ett (under)rum (till R^n) är en mängd H i R^n med egenskaperna:

- Nollvektorn är i H .
- Om u och v är i H , så är $u+v$ i H .
- Om u är i H och c är skalär, så är cu i H .

En bas för rummet H är en mängd linjärt oberoende vektorer som spänner upp H , dvs alla vektorer i H kan uttryckas som en unik linjärkombination av basvektorerna.

2.8 Vektor-rum

Col A: alla linjärkombinationer av kolonner i A

Bas: A~B, kolla pivotkolonner i B, ta motsvarande kolonner i A

$Ax=0$ och $Bx=0$ har samma lösningar,

dvs kolumner i A och kolumner i B har samma linjära beroenden.

Därför kan vi identifiera pivotkolumner i B och ta motsvarande kolumner i A som bas för Col A.

Row A: alla linjärkombinationer av rader i A

Bas: A~B, B trappstegsform, ickenollrader ur B

Row A = Row B och rader i B är lin. ober. så de funkar som bas.

Nul A: alla lösningar till $Ax=0$ (nollrum till $T(x)=Ax$ kallas kernel of T)

Bas: varje parameter i lösning ger en basvektor

Col A och Row A är uppenbart rum.

Nul A:

$A0=0$, så 0 ligger i Nul A.

$Acu=cAu=c0=0$, så cu ligger i Nul A om u gör det.

$A(u+v)=Au+Av=0+0=0$, så $u+v$ ligger i Nul A om u och v gör det.



Fraunhofer

CHALMERS

Research Centre

Industrial Mathematics

2.9 Koordinatvektor

Vektorn x 's koordinater i basen $B = \{b_1, \dots, b_p\}$ kan skrivas

unikt som koordinatvektorn $[x]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix}$

$$\text{dvs} \quad x = c_1 b_1 + \dots + c_p b_p = B \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix} = x$$

Alltså kan vi lösa för $\begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix}$ genom att ställa upp $[B \ x] \sim \dots$

2.9 Dimension av rum

$\text{Dim}(H)$ =antal vektorer i bas för H

$\text{Rang}(A)=\text{dim}(\text{Col}(A))=\text{dim}(\text{Row}(A))$

$\text{Rang}(A)+\text{dim}(\text{Nul}(A))=n$ (A har n kolonner)

$A \sim B$ (radreducerad)

Antal pivotkolumner i A och B = antal ickenollrader i B

Därför är $\text{dim}(\text{Col}(A))=\text{dim}(\text{Row}(A))$ och vi kallar det Rang.

$\text{Rang}(A)$ är antal pivotkolumner.

Varje ickepivotkolumn ger en fri variabel och en basvektor för $\text{Nul}(A)$.

Så $\text{Rang}(A)+\text{dim}(\text{Nul}(A))=n$

2.9 Notera

H p -dimensionellt underrum till R^n .

Alla mängder med mer än p vektorer är linjärt beroende.

Alla lin.ober. mängder med exakt p vektorer är baser för H .

Alla mängder med exakt p vektorer som spänner upp H är baser för H .

$$S = \{v_1, \dots, v_p\} \quad H = \text{span } S$$

Om en vektor v i S är en linjärkombination av resten, så spänns H upp även om v tas bort ur S .

På så vis kan man jobba sig fram till en bas för H .

2.9 Notera

Följande är ekvivalenta för en $n \times n$ -matris A :

- A inverterbar
- A s kolumner är en bas för R^n
- $\text{Col } A = R^n$
- $\text{Dim} (\text{Col } A) = n$
- $\text{Rang}(A) = n$
- $\text{Nul } A = \{0\}$
- $\text{Dim} (\text{Nul } A) = 0$

3 Determinant av (kvadratisk nxn) matris:

$$\det(A)=|A| = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

$$\det([a]) = a$$

$$\det \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = ad - bc$$

Utveckla längs valfri
rad:

$$\det(A) = \sum_{j=1}^n a_{ij} (-1)^{i+j} \det(A_{ij})$$

eller kolumn:

$$\det(A) = \sum_{i=1}^n a_{ij} (-1)^{i+j} \det(A_{ij})$$

A_{ij} är matrisen
man får om man
raderar rad i och
kolumn j från A.

Om A triangulär så $\det(A)$ =produkt av diagonalelement

$$\det(AB) = \det(A) \det(B)$$

$$\det(A + B) \neq \det(A) + \det(B)$$



3 Determinant av (kvadratisk nxn) matris:

Elementära radoperationer (kolonn-op på samma sätt):

- Addera mult av rad till annan rad: ändrar ej det
- Radbyte: byt tecken på det
- Skala rad med k: skala det med k

Betrakta exempel på radoperationer som mult. med elementära matriser:

$$\det(E_{skalning} * A) = \det\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & k \end{bmatrix}\right) * \det(A) = k * \det(A)$$

$$\det(E_{addmult} * A) = \det\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & m \\ 0 & 0 & 1 \end{bmatrix}\right) * \det(A) = 1 * \det(A)$$

$$\det(E_{radbyte} * A) = \det\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\right) * \det(A) = 1 * \det\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right) * \det(A) = 1 * (-1) * \det(A)$$

$$\det(cA) = c^n \det(A)$$

(skala varje rad)

$$\det A = \det A^T$$

(utveckla längs rad i A är samma sak som att utveckla längs kolumn i A^T)



3 Determinant och area/volym:

$\text{Det}(A)$ är lika med arean/volymen av parallelogram/parallelepiped som spänns upp av kolumnerna i A .

Om T är en linjär transformation bestämd av A och S är en ändlig region så är "arean/volymen" av $T(S)$ lika med determinanten av A multiplicerat med arean/volymen av S .

3 Determinant och inverterbarhet

A är inverterbar om $\det(A) \neq 0$

Radreducera till trappstegsform.

För ickeinverterbar A fås nollrad.

Utveckla $\det(A)$ längs nollrad. $\det(A) = 0$

För inverterbar A fås nollskilda pivotelement på diagonalen.

$\det(A) = \text{produkt av diagonalelement} \neq 0$

Finns explicit formel för inversen med hjälp av determinanter, men i praktiken är den metoden vi redan lärt oss oftast bättre.

$$A^{-1} = \frac{\text{adj}A}{\det A} \quad \text{adj}A = \begin{bmatrix} c_{11} & \cdots & c_{n1} \\ \vdots & \ddots & \vdots \\ c_{1n} & \cdots & c_{nn} \end{bmatrix} \quad c_{ij} = (-1)^{i+j} \det A_{ij}$$



3 Cramers regel

A inverterbar

$Ax=b$ har unika lösningen x med $x_i = \frac{\det A_i(b)}{\det A}$

$A_i(b)$ är A med kolumn i ersatt av b .

Oftast är de metoder vi lärt oss förut bättre för att beräkna x ,
men ibland är denna regel användbar.

4.1-4.3 Vektor-rum inte bara i R^n

Kap 4 jämfört med kap 2:

Den enda skillnaden är att vi i kapitel 2 jobbade med R^n .

Nu jobbar vi ett generellt vektorrum, men allt fungerar som vanligt (men generella vektorrum kan vara oändligt-dimensionella).

Definition vektorrum: en mängd V av objekt kallade vektorer med två operationer kallade addition och multiplikation med skalär där följande gäller för alla u, v och w i V och alla skalärer c och d .

1. $u+v$ ligger i V
2. $u+v=v+u$
3. $(u+v)+w=u+(v+w)$
4. Det finns vektor 0 så att $u+0=u$
5. För varje u finns $-u$ så att $u+(-u)=0$
6. cu ligger i V
7. $c(u+v)=cu+cv$
8. $(c+d)u=cu+du$
9. $c(du)=(cd)u$
10. $1u=u$ (mult med skalären 1)



4.4-4.5 Koordinatsystem

x är koordinater för punkten x i standardbasen

$[x]_B$ är koordinater för punkten x i basen B

$P_B = [b_1 \quad \dots \quad b_n]$ är inverterbar eftersom B är en bas.

$$P_B[x]_B = [b_1 \quad \dots \quad b_n] \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = c_1 b_1 + \dots + c_n b_n = x$$

$$x = P_B[x]_B \quad [x]_B = P_B^{-1}x$$

4.4-4.5 Koordinatsystem

$[x]_B$ är unik.

Antag att vi också har $d_1 b_1 + \dots + d_n b_n = x$

$$0 = x - x = (c_1 - d_1)b_1 + \dots$$

Eftersom $b_1 \dots b_n$ är lin. ober. så är alla $c_i - d_i = 0$, dvs $c_i = d_i$

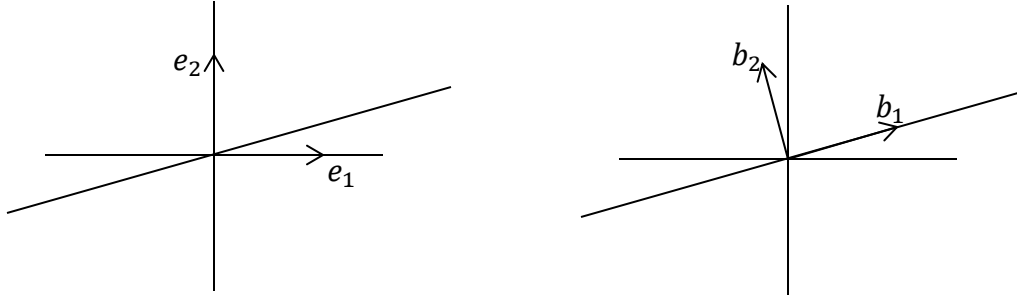
x kan ligga i ett generellt vektorrum V

$[x]_B$ ligger i R^n

$x \rightarrow [x]_B$ är en one-to-one, onto, linjär transformation (isomorfism)

4.6-4.7 Varför byta bas?

Tex för att vissa operationer blir enklare i en speciell bas, tex spegling i linje:



I den nya basen speglas $c_1b_1 + c_2b_2$ till $c_1b_1 - c_2b_2$, dvs
 $T(x) = Ax$ med den enkla matrisen $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

4.6-4.7 Basbyte

$B = \{b_1, \dots, b_n\}$ och $C = \{c_1, \dots, c_n\}$ är två baser för V .

Då finns en unik basbytesmatris $P_{C \leftarrow B}$ så att

$$[x]_C = P_{C \leftarrow B} [x]_B \quad [x]_B = P_{C \leftarrow B}^{-1} [x]_C \quad P_{B \leftarrow C} = P_{C \leftarrow B}^{-1}$$

Tre sätt att beräkna basbytesmatrisen:

$$P_{C \leftarrow B} = P_C^{-1} P_B$$
$$([x]_C = P_C^{-1} x = P_C^{-1} P_B [x]_B)$$

$$P_{C \leftarrow B} = [[b_1]_C \quad \dots \quad [b_n]_C]$$
$$(P_C^{-1} P_B = P_C^{-1} [b_1 \quad \dots \quad b_n] = [P_C^{-1} b_1 \quad \dots \quad P_C^{-1} b_n] = [[b_1]_C \quad \dots \quad [b_n]_C])$$

$$[P_C \ P_B] \sim [I \ P_{C \leftarrow B}]$$
$$(P_C P_{C \leftarrow B} = P_B)$$



4.6-4.7 Basbyte för transformation

A matris för transformation i basen B.

Vad är matrisen för transformationen i basen C?

Tänk så här:

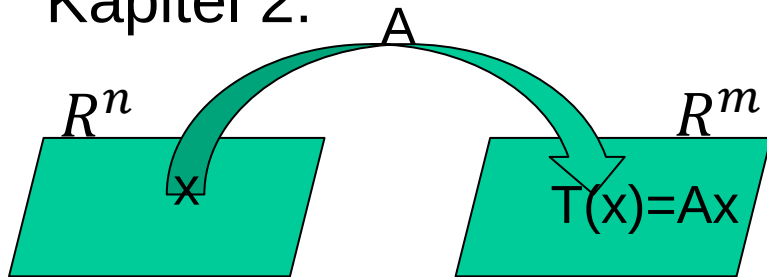
1. Vi startar ju nu med vektorer v i basen C.
2. Byt från C till B, dvs applicera $P_{B \leftarrow C}$, alltså $P_{B \leftarrow C}[v]_C$
3. Applicera transformationen A, alltså $A(P_{B \leftarrow C}[v]_C)$
4. Byt tillbaka från B till C, dvs applicera $P_{B \leftarrow C}^{-1}$, alltså $P_{B \leftarrow C}^{-1}(AP_{B \leftarrow C}[v]_C)$

Transformationsmatrisen i basen C blir alltså $P_{B \leftarrow C}^{-1}AP_{B \leftarrow C}$

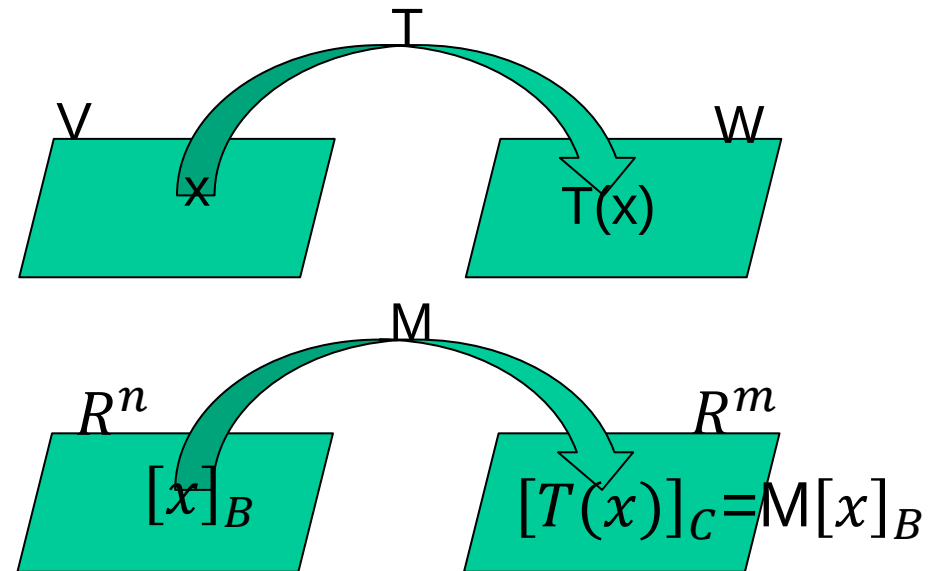


4.7 Hantera generella vektorrum via koordinatvektor

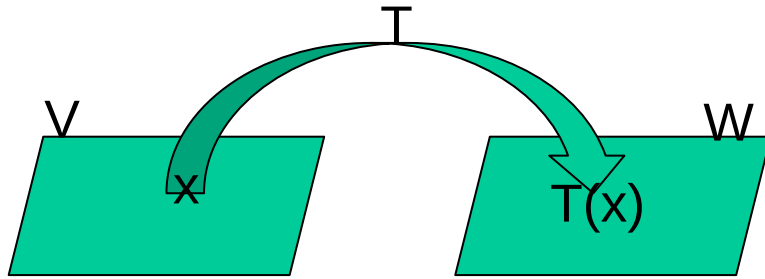
Kapitel 2:



Kapitel 4:

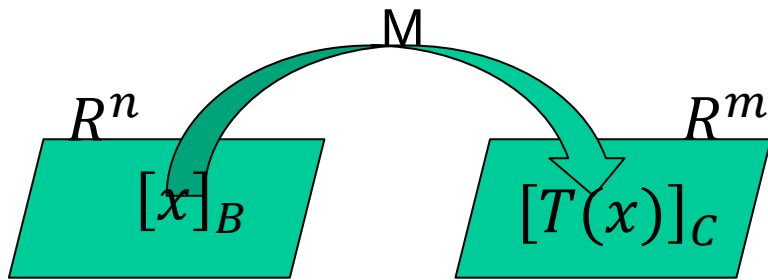


4.7 Matris för transformation av koordinatvektor



B bas för V & C bas för W

$$[T(x)]_C = M[x]_B$$
$$M = [[T(b_1)]_C \quad \dots \quad [T(b_n)]_C]$$



$$\begin{aligned} [T(x)]_C &= P_C^{-1} T(x) \\ &= P_C^{-1} T(P_B[x]_B) \\ &= P_C^{-1} [T(b_1) \quad \dots \quad T(b_n)] [x]_B \\ &= [P_C^{-1} T(b_1) \quad \dots \quad P_C^{-1} T(b_n)] [x]_B \\ &= [[T(b_1)]_C \quad \dots \quad [T(b_n)]_C] [x]_B \end{aligned}$$

Om $V=W$ och $B=C$ så kallas M för B -matrisen för T ; $[T]_B$
Jfr $A = [T(e_1) \dots T(e_n)]$

5 Egenvärden & egenvektorer

$$Ax = \lambda x \quad \lambda \text{ egenvärde} \quad x \text{ egenvektor } (\neq 0)$$

dvs $(A - \lambda I)x = 0$ har icke triviala lösningar

dvs $(A - \lambda I)$ ej inverterbar

$$\text{i.e. } \det(A - \lambda I) = 0$$

5 Beräkning av egenvärden

Detta kallas den karaktäristiska ekvationen, vänsterledet är ett polynom av grad n om A är $n \times n$.

Lösningarna till denna ekvation är egenvärdena.

För $n=2$ finns ju exakt formel. För högre n så får man försöka faktorisera polynomet (Matlab etc löser numeriskt).

Egenvärden kan förekomma med multiplicitet större än 1 (betyder att faktorisering har faktorn $(\lambda - a)$ multipelt).

Det finns alltid exakt n lösningar (om man tillåter komplexa egenvärden och räknar med multiplicitet).

Egenvärdena till triangulär matris är diagonalelementen.

$$\det(A - \lambda I) = \det \begin{pmatrix} a_{11} - \lambda & ? & ? \\ 0 & a_{22} - \lambda & ? \\ 0 & 0 & \ddots \end{pmatrix} = (a_{11} - \lambda)(a_{22} - \lambda)$$



5 Komplexa egenvärden

Även om A är reell, så kan man få komplexa egenvärden och egenvektorer.

De förekommer i så fall i komplexkonjugerade par.

Anta att λ är egenvärde med egenvektor x .

$$A\bar{x} = \overline{Ax} = \overline{\lambda x} = \bar{\lambda}\bar{x} \quad \text{ty}$$

$$\begin{aligned}\overline{(a + bi)(c + di)} &= \overline{(ac - bd) + (bc + ad)i} \\ &= (ac - bd) - (bc + ad)i\end{aligned}$$

$$\begin{aligned}\overline{(a + bi)} * \overline{(c + di)} &= (a - bi)(c - di) \\ &= (ac - bd) - (bc + ad)i\end{aligned}$$

Alltså är även $\bar{\lambda}$ ett egenvärde med egenvektor $\bar{x}$.



5 Egenvärden till inverterbar matris

Något $\lambda=0$ betyder att $Ax=0$ har icke-triviala lösningar, vilket betyder att A inte är inverterbar, Så A inverterbar om inget eget värde är 0.

5 Eigenvektorer

Hitta egenvektorer genom att sätta in egenvärdena λ i

$(A - \lambda I)x = 0$ och lösa för x .

Det blir alltid parameterlösningar till varje λ .

Representera med en bas.

(jfr bas för nollrum till $A - \lambda I$)

Mängden av alla egenvektorer till ett egenvärde utgör ett rum.

Linjärkombinationer av egenvektorer till ett egenvärde är egenvektorer till samma egenvärde.

Mängden av alla egenvektorer till ett egenvärde kan representeras av en bas.

5 Egenvektorsrum & diagonalisering

Sats: A ($n \times n$) diagonaliserbar $A = PDP^{-1}$ med D diagonal
om det finns en bas till $\mathbb{R}^n$ av egenvektorer till A (dvs det finns n lin. ober. egenvektorer till A).

De bildar då kolonner i P och D har motsvarande egenvärden.

Bevis: $P = [p_1 \quad \cdots \quad p_n] \quad D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix} \quad AP = [Ap_1 \quad \cdots \quad Ap_n] \quad PD = [\lambda_1 p_1 \quad \cdots \quad \lambda_n p_n]$

Antag $A = PDP^{-1}$,
 $AP = PD$, $Ap_i = \lambda_i p_i$, dvs egenvärden och egenvektorer, och egenvektorerna är lin.ober
eftersom P är inverterbar.

Antag $p_1 \quad \cdots \quad p_n$ är lin. ober. egenvektorer och $\lambda_1 \quad \cdots \quad \lambda_n$ är motsvarande egenvärden.

$$Ap_i = \lambda_i p_i$$

$$AP = PD$$

P inverterbar eftersom egenvektorerna lin. ober.

$$A = PDP^{-1}$$



5 Olika egenvärden och diagonalisering

Diagonalisering alltid möjligt om $n \times n$ -matrisen A har n olika egenvärden,
ty egenvektorer från olika egenvärden är linjärt oberoende
(enligt sats vars bevis vi inte går genom).

Om A har p olika egenvärden:

Dimensionen av egenvektorsrummet för ett egenvärde är mindre än eller lika med multipliciteten av egenvärdet.

A diagonaliserbar om dimensionen är lika med multipliciteten för alla egenvärden.

5 (Sats 8 i 5.4)

$A = PDP^{-1}$ B är basen av kolumner i P.

Då är D B-matrisen för transformationen $T(x)=Ax$,

Dvs i basen B representeras transformationen enkelt med en diagonalmatris.

$$\begin{aligned}[T]_B &= [[T(b_1)]_B \quad \dots \quad [T(b_n)]_B] = [[Ab_1]_B \quad \dots \quad [Ab_n]_B] = \\ &[P^{-1}Ab_1 \quad \dots \quad P^{-1}Ab_n] = P^{-1}A[b_1 \quad \dots \quad b_n] = P^{-1}AP = \\ &P^{-1}(PDP^{-1})P = D\end{aligned}$$

5 Determinanter, egenvärden & likformiga matriser

$\det(A)$ = produkten av egenvärdena till A

Kom ihåg att för att beräkna egenvärden gör vi så här:

$$\det(A - \lambda I) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$$

Men denna ekvation gäller ju för alla λ så vi kan sätta $\lambda = 0$ vilket ger $\det(A) = \lambda_1 \cdots \lambda_n$

A och B är likformiga (similar) om $P^{-1}AP = B$ för något P .
Likformiga matriser har samma egenvärden.

$$\begin{aligned}\det(B - \lambda I) &= \det(P^{-1}AP - \lambda P^{-1}P) = \det(P^{-1}(A - \lambda I)P) \\ &= \det(P^{-1}) \det(A - \lambda I) \det(P) \\ &= \det(A - \lambda I) \det(P^{-1}) \det(P) \\ &= \det(A - \lambda I) \det(P^{-1}P) = \det(A - \lambda I) \det(I) \\ &= \det(A - \lambda I)\end{aligned}$$

5 Dynamiska system

$$x_{k+1} = Ax_k$$

Eventuell jämvikt antas i egenvektorn $\hat{x}$ till egenvärdet 1.

Om x tex beskriver en fördelning på olika tillstånd så konvergerar systemet mot denna jämvikt.

Antag n lin. ober. egenvektorer i R^n .

Egenvektorerna är då en bas för R^n och alla x i R^n kan skrivas

$$x_0 = c_1 v_1 + \dots + c_n v_n \quad (v \text{ egenvektorer})$$

$$x_1 = Ax_0 = c_1 \lambda_1 v_1 + \dots + c_n \lambda_n v_n$$

$$x_k = c_1 (\lambda_1)^k v_1 + \dots + c_n (\lambda_n)^k v_n$$

Om k får gå mot oändligheten så kommer det egenvärde med störst absolutbelopp att dominera fullständigt (om inte motsvarande $c=0$)



5 Differentialekvationer

$$x' = Ax \quad x_0 \text{ initialvärde för } x$$

v_i egenvektorer till A med motsvarande egenvärden λ_i

$$x(t) = c_1 v_1 e^{\lambda_1 t} + \dots + c_n v_n e^{\lambda_n t}$$
$$\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = P^{-1} x_0$$

5 Bevis

Variabelbytet $y = P^{-1}x$ ger det enklare systemet $\begin{cases} y'_1 = \lambda_1 y_1 \\ y'_2 = \lambda_2 y_2 \\ \vdots \end{cases}$

Vilket har lösningen $y = \begin{bmatrix} c_1 e^{\lambda_1 t} \\ \vdots \end{bmatrix} \quad \begin{bmatrix} c_1 \\ \vdots \end{bmatrix} = y(0) = P^{-1}x_0$

Vilket ger $x = Py = [v_1 \quad \cdots]y = c_1 v_1 e^{\lambda_1 t} + \dots$

6.1 Skalarprodukt / inner product / dot product

Skalarprodukt (i R^n).

u och v kolonnvektorer

$$u \cdot v = u^T v$$

$$\|v\| = \sqrt{v \cdot v}$$

u , v och w vektorer i R^n och c skalär

- $u \cdot v = v \cdot u$
- $(u + v) \cdot w = u \cdot w + v \cdot w$
- $(cu) \cdot v = c(u \cdot v) = u \cdot (cv)$
- $u \cdot u \geq 0$ ($u \cdot u = 0$ om $u = 0$)

Normalisering: $u = \frac{v}{\|v\|}$

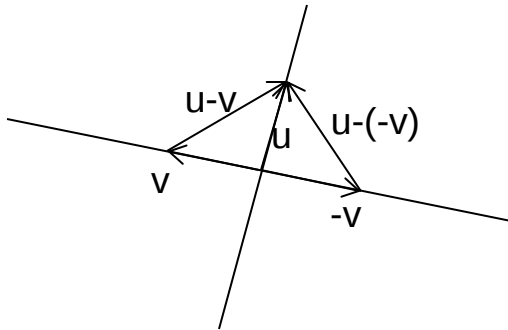
$$u \cdot v = \|u\| * \|v\| * \cos(\text{vinkeln mellan } u \text{ och } v)$$

Avstånd mellan två punkter är normen (längden) på vektorn mellan dem.



6.1 Ortogonalitet

Två linjer genom origo definierade av u och v .



Linjerna är ortogonala (vinkelräta) omm

$$\|u - v\| = \|u - (-v)\| \Leftrightarrow \|u - v\|^2 = \|u - (-v)\|^2$$

$$\|u - v\|^2 = (u - v) \cdot (u - v) = u \cdot u + v \cdot v - 2u \cdot v$$

$$\|u - (-v)\|^2 = (u + v) \cdot (u + v) = u \cdot u + v \cdot v + 2u \cdot v$$

$$\text{Dvs ortogonalt omm } 2u \cdot v = -2u \cdot v \Leftrightarrow u \cdot v = 0$$



6.1 Ortogonalitet

u ortogonal mot v omm $u \cdot v = 0$

u ortogonal mot W omm $u \cdot v = 0$ för alla v i W

u ortogonal mot v omm $\|u + v\|^2 = \|u\|^2 + \|v\|^2$

Mängden av vektorer ortogonala mot W kallas ortogonala komplementet till W .

x är ortogonal mot W omm x är ortogonal mot varje vektor i en mängd som spänner upp W .

En mängd vektorer kallas ortogonal om varje par av vektorer i mängden är ortogonala.

Vektorerna i en ortogonal mängd är lin. ober.

$$c_1 v_1 + \cdots + c_n v_n = 0$$

$$\begin{aligned} 0 = 0v_i &= (c_1 v_1 + \cdots + c_n v_n) \cdot v_i = c_1 (v_1 \cdot v_i) + \cdots + c_n (v_n \cdot v_i) = c_i (v_i \cdot v_i) \\ &= c_i a \quad a \neq 0 \quad \Rightarrow c_i = 0 \text{ för alla } i \end{aligned}$$



6.1 ONbas

En ortogonal mängd kallas ortonormerad (ortonormal) om vektorerna är normerade till enhetsvektorer (längd 1).

En ortogonal bas är en bas som även är en ortogonal mängd. En ortonormerad bas (ONbas) har dessutom normerade basvektorer (längd 1).

6.2 Ortogonal matris

U har ortonormerade (ortogonala med längd 1) kolonner (& rader) (kallas att U är ortogonal) omm $U^T U = I$ dvs $U^T = U^{-1}$

$$U = [u_1 \quad \cdots \quad u_n]$$
$$U^T U = \begin{bmatrix} u_1^T u_1 & u_1^T u_2 & \cdots \\ u_2^T u_1 & u_2^T u_2 & \\ \vdots & & \ddots \end{bmatrix}$$

Dvs $U^T U = I$ omm $u_i^T u_i = 1$ och $u_i^T u_j = 0$ dvs ortonormerade.

Om U ortogonal så:

a) $\|Ux\| = \|x\|$

b) $(Ux) \cdot (Uy) = x \cdot y$

c) $(Ux) \cdot (Uy) = 0$ omm $x \cdot y = 0$

$b \Rightarrow c$) *trivialt*

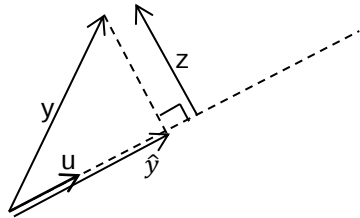
$$b \Rightarrow a) \|Ux\|^2 = (Ux) \cdot (Ux) = x \cdot x = \|x\|^2$$

$$b)(Ux) \cdot (Uy) = (x_1 u_1 + \cdots + x_n u_n) \cdot (y_1 u_1 + \cdots + y_n u_n) = x_1 y_1 + \cdots x_n y_n \\ = x \cdot y$$



6.2 Projektion på linje definierad av vektor

Ortogonal projektion av y på u : $\hat{y} = \frac{y \cdot u}{u \cdot u} u$



$$\hat{y} = a * u$$

$$y = \hat{y} + z \Leftrightarrow z = y - \hat{y} = y - a * u$$

z ortogonal mot $u \Leftrightarrow$

$$0 = z \cdot u = (y - a * u) \cdot u = y \cdot u - a * u \cdot u \Rightarrow a = \frac{y \cdot u}{u \cdot u} \Rightarrow \hat{y} = \frac{y \cdot u}{u \cdot u} u$$

Om u är normaliserade till längd 1:

$$\hat{y} = \frac{y \cdot u}{u \cdot u} u = y \cdot u * u \quad \text{där } y \cdot u \text{ är längden och } u \text{ riktningen}$$

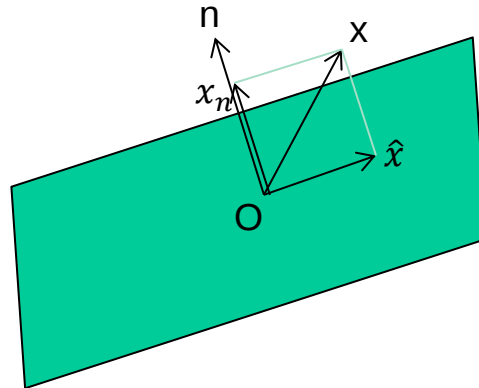
6.2 Projektion på plan genom origo

Plan: Alla punkter x (vektorer från origo) som uppfyller
 $ax_1 + bx_2 + cx_3 = 0$

Normal (vektor ortogonal mot planet): (a,b,c)

Projektion av godtycklig vektor x på normalen n : $x_n = \frac{x \cdot n}{n \cdot n} n$

Projektion av x på planet: $\hat{x} = x - x_n$



6.3 Projektion av vektorn y till rum W med ortogonal bas B

y kan ligga i eller utanför W .

Observera att metoden inte fungerar om B inte är ortogonal!

$$\begin{aligned} 0 &= (y - \text{proj}_W(y)) \cdot b_i = (y - c_1 b_1 - \cdots - c_n b_n) \cdot b_i \\ &= y \cdot b_i - c_i (b_i \cdot b_i) \end{aligned}$$

$$c_i = \frac{y \cdot b_i}{b_i \cdot b_i} \quad \text{proj}_W(y) = \left(\frac{y \cdot b_1}{b_1 \cdot b_1} \right) b_1 + \cdots + \left(\frac{y \cdot b_p}{b_p \cdot b_p} \right) b_p$$

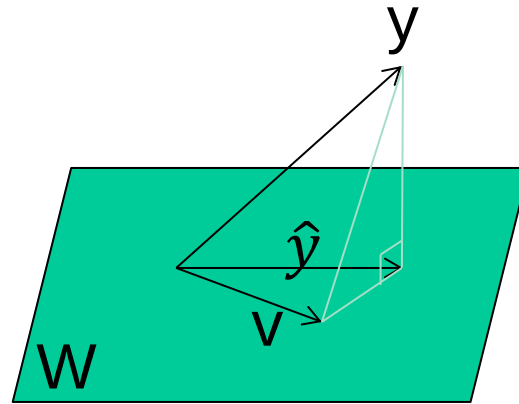
Med ortonormerad bas (ONbas) B för W :

$$c_i = y \cdot b_i \quad \text{proj}_W(y) = (y \cdot b_1) b_1 + \cdots + (y \cdot b_p) b_p = BB^T y$$

6.3 Närmaste punkt

$\hat{y} = \text{proj}_W(y)$ är den punkt i W som är närmast y

$\|y - \hat{y}\| < \|y - v\|$ för alla v i W $v \neq \hat{y}$



$y - \hat{y}$ är ortogonal mot $v - \hat{y}$

Pythagoras sats: $\|y - \hat{y}\|^2 + \|v - \hat{y}\|^2 = \|y - v\|^2$

$\|v - \hat{y}\|^2 > 0$ eftersom $v \neq \hat{y}$

Alltså: $\|y - \hat{y}\|^2 < \|y - v\|^2$

6.4 Skapa en ortonormerad (ON) bas

Gram-Schmidt:

Skapa ONbas $\{v_1 \dots v_p\}$ från bas $\{x_1 \dots x_p\}$:

$$v_1 = x_1$$

$$v_2 = x_2 - \frac{x_2 v_1}{v_1 v_1} v_1$$

$$v_3 = x_3 - \frac{x_3 v_1}{v_1 v_1} v_1 - \frac{x_3 v_2}{v_2 v_2} v_2$$

etc.

Normalisera i varje steg eller för alla efteråt

6.4 QR-faktorisering

A en $m \times n$ -matris med lin. ober. kolumner

$$A = QR$$

Q $m \times n$ vars kolumner är en ONbas för $\text{Col} A$

R $n \times n$ övretriangulär inverterbar med positiva diagonalelement

Q fås tex via Gram-Schmidt.

R fås antingen också via Gram-Schmidt-processen, eller via

$$Q^T A = Q^T (QR) = IR = R$$

6.4 QR-faktorisering

Bevis:

Kolumner i A är en bas för $\text{Col}A$ (ty lin.ober.). $A = [x_1 \ \cdots \ x_n]$

Med tex Gram-Schmidt kan man då få en ONbas Q för $\text{Col}A$.

$$Q = [u_1 \ \cdots \ u_n]$$

Gram-Schmidt-processen medför att $x_k \in \text{span}\{u_1 \ \cdots \ u_k\}$

dvs $x_k = r_{1k}u_1 + \cdots + r_{kk}u_k$ (om $r_{kk} < 0$ så multiplicera r_{kk} och u_k med -1)

$$r_k = \begin{bmatrix} r_{1k} \\ \vdots \\ r_{kk} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad x_k = Qr_k \quad R = [r_1 \ \cdots \ r_n] \quad A = [Qr_1 \ \cdots \ Qr_n] = QR$$

R kvadratisk så inverterbar om lin.ober. kolumner.

$Rx = 0 \Leftrightarrow QRx = Q0 = 0 \Leftrightarrow Ax = 0$ och eftersom A har lin.ober. kol. så $x = 0$

Alltså har R lin.ober.kol. och är därmed inverterbar.

Att R är inverterbar och triangulär medför att den har nollskilda diagonalelement.



6.5-6.6 Approximativ Minsta-kvadrat-lösning

$Ax=b$ har inte alltid någon exakt lösning

(i praktiken finns nästan aldrig någon exakt lösning).

Då kan man istället bestämma den lösning som är så nära att uppfylla ekvationen som möjligt.

Om "nära" mäts som $\|Ax - b\|$ är så liten som möjligt så kallas detta minstakvadratlösning.

6.5-6.6 Approximativ Minsta-kvadrat-lösning

Minstakvadratlösningen till $Ax = b$ är lösningen till $A^T Ax = A^T b$

Ax är $\text{Col}A$ om x varieras fritt.

Den närmaste punkten i $\text{Col}A$ till b är $\hat{b} = \text{proj}_{\text{Col}A}(b)$

Så vi söker $\hat{x}$ i $A\hat{x} = \hat{b}$ och $b - \hat{b}$ är ortogonal mot $\text{Col}A$, dvs $b - A\hat{x}$ är ortogonal mot $\text{Col}A$, dvs ortogonal mot varje kolumn a_j i A , dvs $a_j^T(b - A\hat{x}) = 0$, eller med alla kolumner på en gång: $A^T(b - A\hat{x}) = 0 \Leftrightarrow A^T b - A^T A\hat{x} =$

$$0 \Leftrightarrow A^T A\hat{x} = A^T b$$

Dvs $\hat{x}$ uppfyller de så kallade normalekvationerna.

Andra hållet: Antag att $\hat{x}$ uppfyller $A^T A\hat{x} = A^T b$

Då är alltså $b - A\hat{x}$ ortogonal mot $\text{Col}A$, och vi kan skriva

$b = A\hat{x} + (b - A\hat{x})$ där $A\hat{x}$ ligger i $\text{Col}A$ och därmed

$$A\hat{x} = \text{proj}_{\text{Col}A}(b) = \hat{b}$$

Dvs $\hat{x}$ är en lösning.



6.5-6.6 Approximativ Minsta-kvadrat-lösning

Om kolumnerna i A är linjärt oberoende så är $A^T A$ inverterbar.

Då finns en unik minstakvadratlösning $\hat{x} = (A^T A)^{-1} A^T b$

$\|A\hat{x} - b\|$ kallas minstakvadratfel

6.5-6.6 Minsta-kvadrat-lösning via QR

$$A = QR$$
$$\hat{x} = R^{-1}Q^T b$$

$A\hat{x} = QR\hat{x} = QRR^{-1}Q^T b = QQ^T b = \text{proj}_{\text{col}A}(b) = \hat{b}$
så $\hat{x}$ ovan är minstakvadratlösningen.

I praktiken inverterar man inte R utan löser $Rx = Q^T b$ vilket är enkelt eftersom R är triangulär.

6.6 Anpassning av kurva till datapunkter

För punkter på linje gäller $y = \beta_0 + \beta_1 x = [1 \quad x] \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$

Med flera data punkter: $y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = X\beta$

I praktiken finns ingen exakt lösning så vi kör
minstakvadratlösning

Dvs lös $X^T X \beta = X^T y$

(för andra kurvor än linjer och för ytor istället för kurvor i fler dimensioner blir X annorlunda med det funkar likadant)

7.1 Diagonalisering av symmetriska matriser

A ortogonalt diagonaliserbar om $A = PDP^T$ D diagonal, P ortogonal

Om $A = A^T$ så kallas A symmetrisk

Om A är symmetrisk så är

- egenvektorer för olika egenvärden inte bara linjärt oberoende utan ortogonala.
- Dimensionen på egenvektorsrum = multiplicitet av egenvärde

A är ortogonalt diagonaliserbar om A är symmetrisk.

7.1 egenvektorer för olika egenvärden till symmetrisk matris är ortogonala

v_1 och v_2 egenvektorer till olika egenvärden λ_1 och λ_2

$$\lambda_1 v_1 \cdot v_2 = (Av_1)^T v_2 = v_1^T A^T v_2 = v_1^T A v_2 = v_1^T \lambda_2 v_2 = \lambda_2 v_1 \cdot v_2$$

Dvs $(\lambda_1 - \lambda_2)v_1 \cdot v_2 = 0$

Men $\lambda_1 - \lambda_2 \neq 0$ så $v_1 \cdot v_2 = 0$

7.1 Spektral dekomposition

Mängden av egenvärdena kallas spektrum

Spektral dekomposition:

$$A = PDP^T = [u_1 \cdots u_n] \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} u_1^T \\ \vdots \\ u_n^T \end{bmatrix} = \lambda_1 u_1 u_1^T + \cdots + \lambda_n u_n u_n^T$$

7.2 Kvadratiska former

Kvadratisk form: $Q(x) = x^T A x$ A symmetrisk

Det finns ett ortogonalt variabelbyte $x=Py$ som transformerar $x^T A x$ till $y^T D y$ med D diagonal (med egenvärden till A på diagonalen).

Kolumnerna i P kallas principal-axlarna och y är koordinatvektorn för x med dessa som bas.

Eftersom A är symmetrisk kan P väljas ortogonal så att $P^T A P = D$ med D diagonal

$$x = Py \quad x^T A x = (Py)^T A (Py) = y^T P^T A P y = y^T D y$$

$$y^T D y \text{ betyder inga blandade termer } y_i y_j \quad i \neq j \quad y^T D y = \sum \lambda_i y_i^2$$

Q (och A) kallas

- Positivt definit om $Q(x) > 0$ för alla x skilda från 0 (semidefinit om $\geq$)
- Negativt definit om $Q(x) < 0$ för alla x skilda från 0 (semidefinit om $\leq$)
- Indefinit om $Q(x)$ antar både positiva och negativa värden.

Q är pos def omm alla egenvärden till A är positiva

Q är neg def omm alla egenvärden till A är negativa

Q är indef omm A har både positiva och negativa egenvärden



7.2-7.3 Mer om kvadratiska former

$$\max_{\|x\|=1} Q(x) = \max\{\lambda_1, \dots, \lambda_n\}$$

$$\min_{\|x\|=1} Q(x) = \min\{\lambda_1, \dots, \lambda_n\}$$

För alla $m \times n$ -matriser A är $A^T A$ symmetrisk och pos. semidef.

Om A och B är sym.pos.semidef. så är $A+B$ sym.pos.semidef.

Skapa symmetriskt A från kvadratisk form genom att dela upp blandtermer i två lika delar.

Tex $7x_1^2 - 6x_1x_2 - 2x_2^2$ ger $A = \begin{bmatrix} 7 & -3 \\ -3 & -2 \end{bmatrix}$