Lesson 11

Or how to find a near solution of inconsistent systems

Chapter 6
Orthogonality and least squares

OPGAVE 2 (fra lektion 10)

Matricen A her er tæt på at være ortogonal

- Hvad vil det sige, at en matrix er ortogonal?
- ”Reparer” på A, således at den bliver ortogonal
  hint: kig på elementet med indeks 3,2

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Søjlevektorerne (og rækkevektorerne) er indbyrdes ortogonale – MEN:

Definition:
Søjler og rækker skal også være *normale* for at vi siger, at matricen A er ortogonal.

$$A = \begin{bmatrix} 0.707 & 0.707 & 0 \\ 0.707 & -0.707 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Inner product (dot product): $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$

$$\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^{\boldsymbol{T}} \boldsymbol{v} = [u_1 \; u_2 \; \cdots \; u_n] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \sum_{i=1}^{n} u_i v_i \in \mathbb{R}$$

- $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$
- $(\boldsymbol{u} + \boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{u} \cdot \boldsymbol{w} + \boldsymbol{v} \cdot \boldsymbol{w}$
- $(c\boldsymbol{u}) \cdot \boldsymbol{v} = \boldsymbol{u} \cdot (c\boldsymbol{v})$
- $\boldsymbol{u} \cdot \boldsymbol{u} \geq 0$ and $\boldsymbol{u} \cdot \boldsymbol{u} = 0 \Leftrightarrow \boldsymbol{u} = 0$
- The **norm** (or length) of a vector is defined as: $\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}} = \sqrt{v_1^2 + \cdots + v_n^2}$
- A vector $\boldsymbol{v}$ is **normalized** (unit vector) if: $\|\boldsymbol{v}\| = 1$
- The **distance** between two vectors is defined as:

$$dist(\boldsymbol{u}, \boldsymbol{v}) = \|\boldsymbol{u} - \boldsymbol{v}\| = \sqrt{(u_1 - v_1)^{\mathbf{2}} + \cdots + (u_n - v_n)^{\mathbf{2}}}$$

- The **angle** θ between two vectors is given by: $\boldsymbol{u} \cdot \boldsymbol{v} = \|\boldsymbol{u}\| \|\boldsymbol{v}\| \cos(\theta)$

Definitions

Orthogonal set

A set of vectors $S = \{\boldsymbol{u}_1 \ \cdots \ \boldsymbol{u}_n\}$ in $\mathbb{R}^n$ with $\boldsymbol{u}_i \perp \boldsymbol{u}_j$ for all $i \neq j$

That is: $\boldsymbol{u}_i \cdot \boldsymbol{u}_j = 0$ for all $i \neq j$

Orthonormal set

An orthogonal set of vectors $S = \{\boldsymbol{u}_1 \ \cdots \ \boldsymbol{u}_n\}$ in $\mathbb{R}^n$ with $\|\boldsymbol{u}_i\| = 1$ for all $\mathrm{i} = 1, \dots, n$.

That is: $\boldsymbol{u}_i \cdot \boldsymbol{u}_j = \delta_{ij} = \begin{cases} 1 & if \ \ i = j \\ 0 & if \ \ i \neq j \end{cases}$

Orthogonal complement $W^{\perp}$

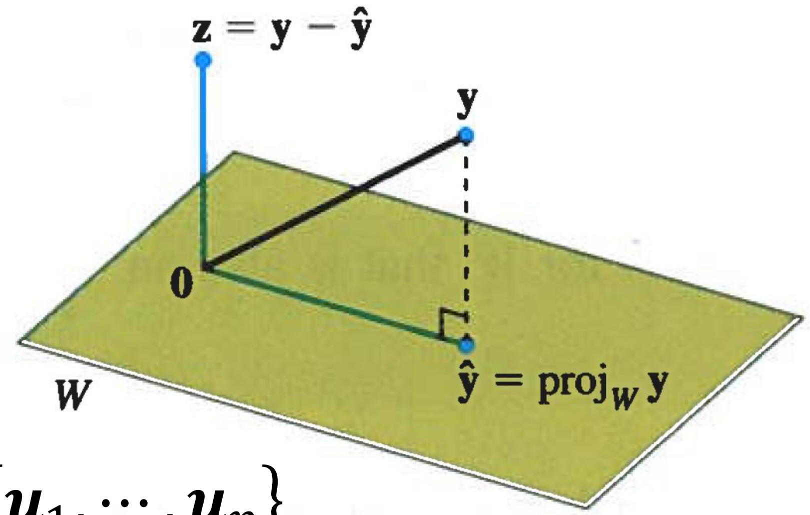
The set of all vectors in $\mathbb{R}^n$ which are orthogonal to all vectors in W.

Theorem 6.8 - The Orthogonal Decomposition Theorem

Let $\{\boldsymbol{u}_1 \ \cdots \ \boldsymbol{u}_n\}$ be an orthogonal basis of $\mathbb{R}^n$.

For each $\boldsymbol{y}$ in $\mathbb{R}^n$:

$$\boldsymbol{y} = c_1\boldsymbol{u}_1 + \cdots + c_n\boldsymbol{u}_n = \widehat{\boldsymbol{y}} + \boldsymbol{z}$$



where

$$\widehat{\boldsymbol{y}} = \overbrace{\frac{\boldsymbol{y}\cdot\boldsymbol{u}_1}{\boldsymbol{u}_1\cdot\boldsymbol{u}_1}}^{c_1}\boldsymbol{u}_1 + \cdots + \overbrace{\frac{\boldsymbol{y}\cdot\boldsymbol{u}_p}{\boldsymbol{u}_p\cdot\boldsymbol{u}_p}}^{c_p}\boldsymbol{u}_p = proj_W\boldsymbol{y} \ \in \ W = Span\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p\}$$

$$\boldsymbol{z} = \boldsymbol{y} - \widehat{\boldsymbol{y}} = \underbrace{\frac{\boldsymbol{y}\cdot\boldsymbol{u}_{p+1}}{\boldsymbol{u}_{p+1}\cdot\boldsymbol{u}_{p+1}}}_{c_{p+1}}\boldsymbol{u}_{p+1} + \cdots + \underbrace{\frac{\boldsymbol{y}\cdot\boldsymbol{u}_n}{\boldsymbol{u}_n\cdot\boldsymbol{u}_n}}_{c_n}\boldsymbol{u}_n = proj_{W^\perp}\boldsymbol{y} \ \in W^\perp = Span\{\boldsymbol{u}_{p+1}, \cdots, \boldsymbol{u}_n\}$$

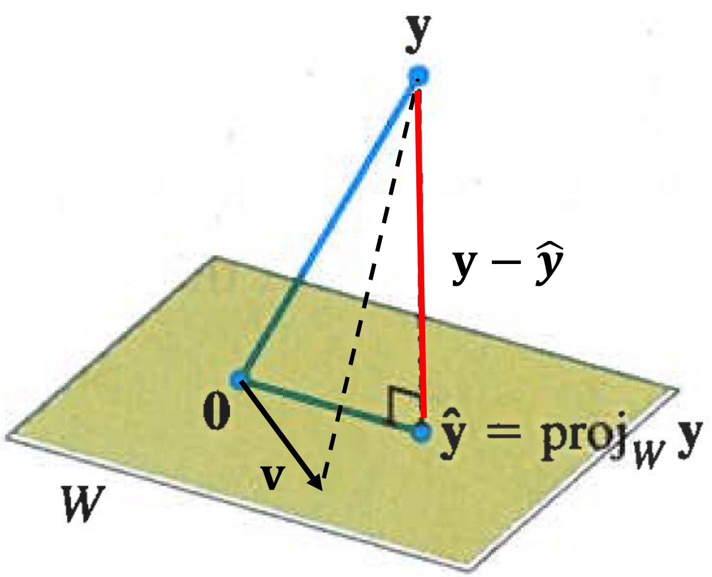
$$\boldsymbol{z} = \boldsymbol{y} - \widehat{\boldsymbol{y}} \perp \widehat{\boldsymbol{y}}$$

Theorem 6.9: The Best Approximation Theorem

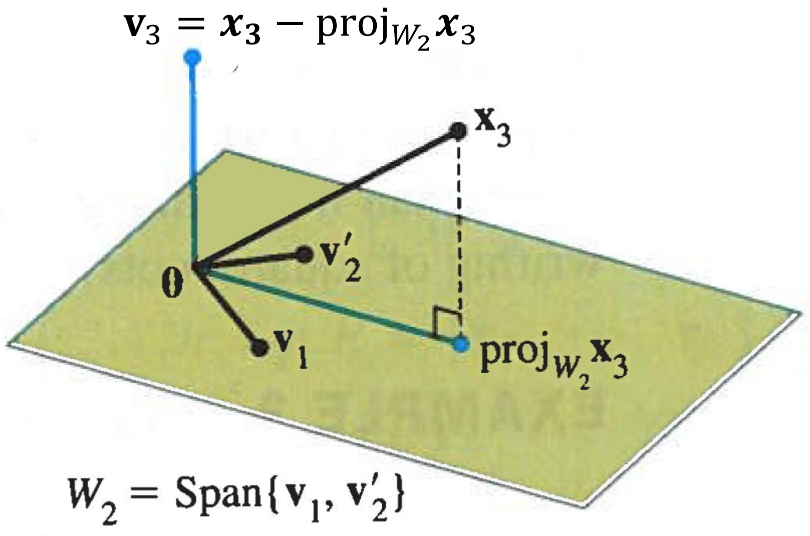
Let W be a subspace of $\mathbb{R}^n$, let $\mathbf{y}$ be any vector in $\mathbb{R}^n$, and let $\hat{\mathbf{y}}$ be the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the closest point in W to $\mathbf{y}$ in the sense that

$$||\mathbf{y} - \hat{\mathbf{y}}|| < ||\mathbf{y} - \mathbf{v}||$$

for all $\mathbf{v}$ in W distinct from $\hat{\mathbf{y}}$.

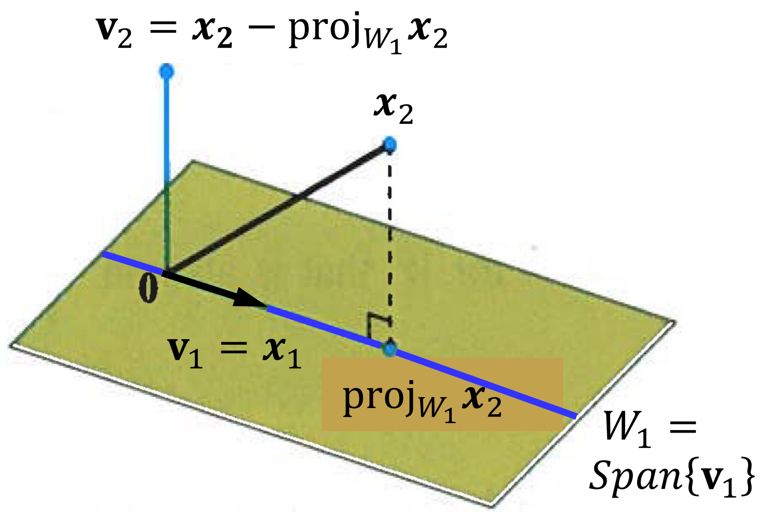


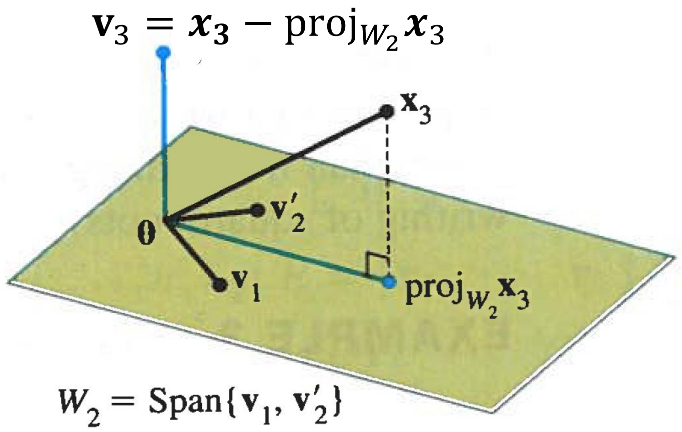
6.4 The Gram-Schmidt Process



Ex 1

$$\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \boldsymbol{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \boldsymbol{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$





The Gram-Schmidt Proces

Given a basis $\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_p\}$ for a subspace W of $\mathbb{R}^n$. Define:

$$\boldsymbol{v}_1 = \boldsymbol{x}_1$$

$$\boldsymbol{v}_2 = \boldsymbol{x}_2 - \frac{\boldsymbol{x_2} \cdot \boldsymbol{v}_1}{\boldsymbol{v_1} \cdot \boldsymbol{v}_1}\boldsymbol{v}_1$$

$$\boldsymbol{v}_3 = \boldsymbol{x}_3 - \frac{\boldsymbol{x_3} \cdot \boldsymbol{v}_1}{\boldsymbol{v_1} \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 - \frac{\boldsymbol{x_3} \cdot \boldsymbol{v}_2}{\boldsymbol{v_2} \cdot \boldsymbol{v}_2}\boldsymbol{v}_2$$

$$\vdots$$

$$\boldsymbol{v}_p = \boldsymbol{x}_p - \frac{\boldsymbol{x_p} \cdot \boldsymbol{v}_1}{\boldsymbol{v_1} \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 - \frac{\boldsymbol{x_p} \cdot \boldsymbol{v}_2}{\boldsymbol{v_2} \cdot \boldsymbol{v}_2}\boldsymbol{v}_2 - \cdots - \frac{\boldsymbol{x_p} \cdot \boldsymbol{v}_{p-1}}{\boldsymbol{v_{p-1}} \cdot \boldsymbol{v}_{p-1}}\boldsymbol{v}_{p-1}$$

Then $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p\}$ is an orthogonal basis for W,

and $Span\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k\} = Span\{\boldsymbol{v}_1, \cdots, \boldsymbol{v_k}\}$ for $k = 1, \cdots, p$

Theorem 6.12: The QR Factorization

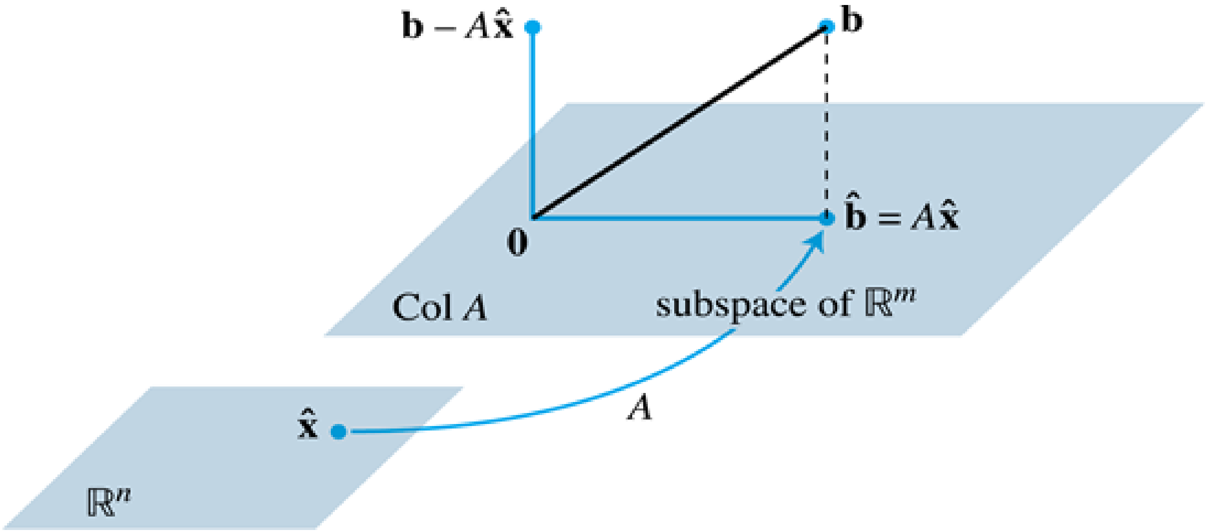
If A is an $m \times n$ matrix with linearly independent columns, then A can be factored as:

$$A = QR$$

where

- Q is an $m \times n$ matrix whose columns form an orthonormal basis for $Col\ A$
- R is an $n \times n$ upper triangle invertible matrix with positive entries on its diagonal.

6.5 Least Squares Problems



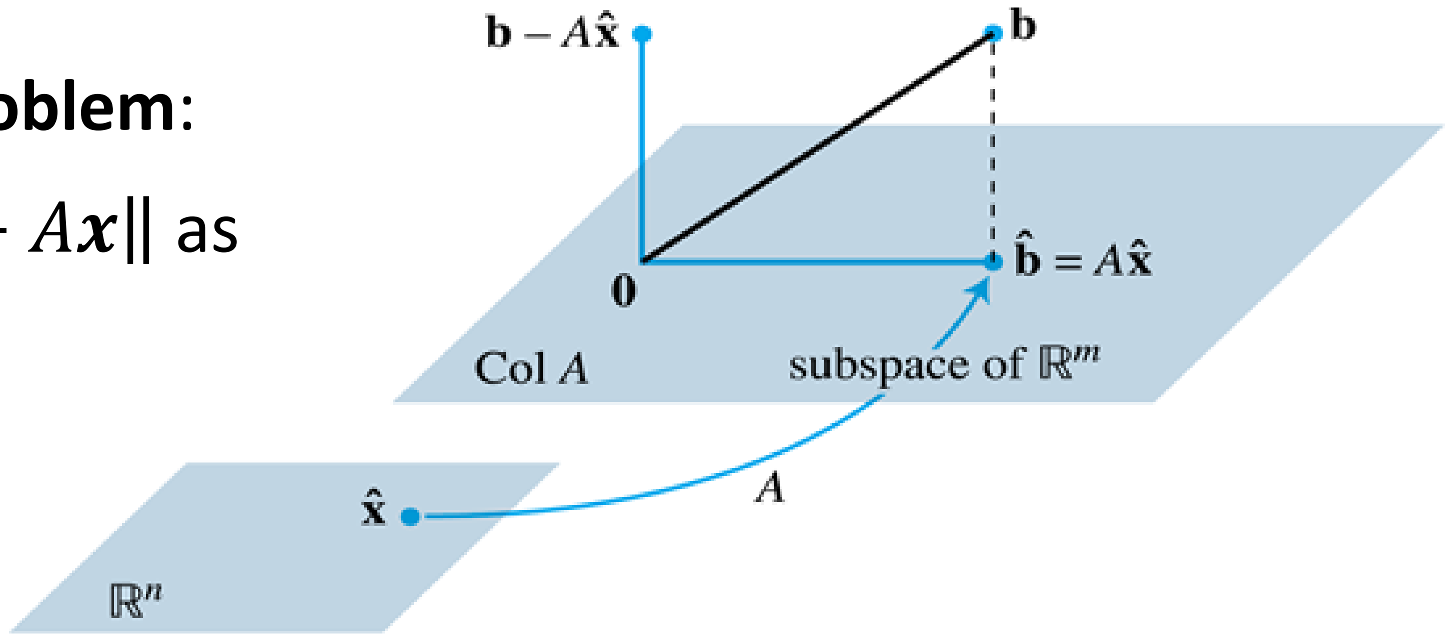
Least-squares problems

Often – in real life – $A\boldsymbol{x} = \boldsymbol{b}$ has no solutions (are inconsistent).
That is, $\boldsymbol{b} \notin Col\ A$.

The best approximation of $\boldsymbol{b}$ is the vector in $Col\ A$, that is closest to $\boldsymbol{b}$.

The general least-squares problem:

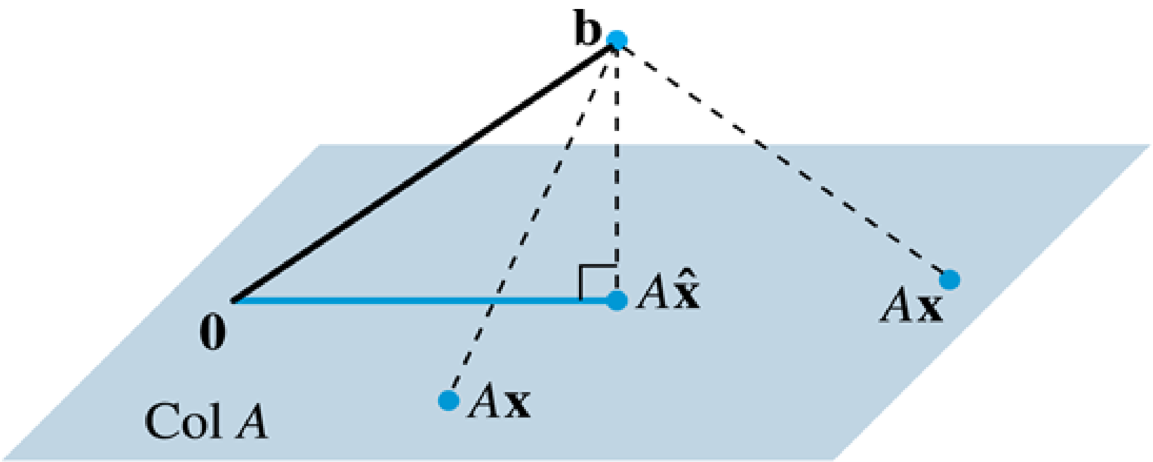
➢ Find an $\boldsymbol{x}$ that makes $||\boldsymbol{b} - A\boldsymbol{x}||$ as small as possible,



Definition:

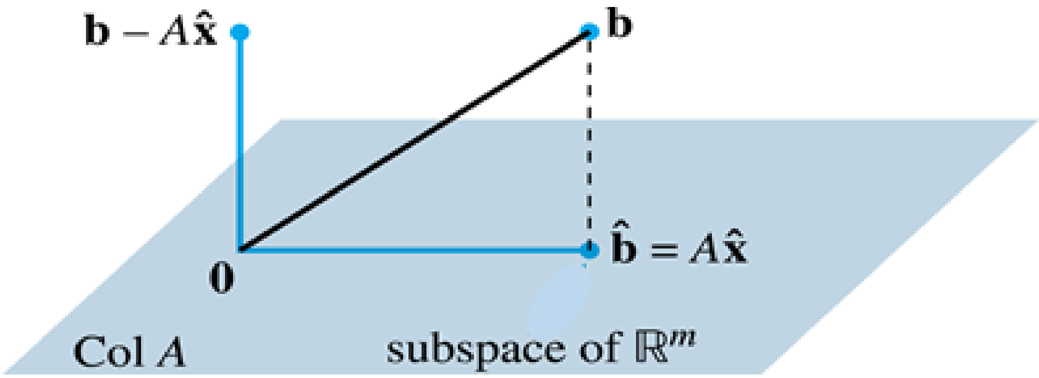
Let A be an $m \times n$ matrix and let $\mathbf{b}$ be in $\mathbb{R}^m$. A least squares solution of $A\mathbf{x} = \mathbf{b}$ is then $\hat{\mathbf{x}}$ in $\mathbb{R}^n$ such that

$$||\mathbf{b} - A\hat{\mathbf{x}}|| \leq ||\mathbf{b} - A\mathbf{x}||, \quad \forall \mathbf{x} \in \mathbb{R}^n$$



By Theorem 6.9:

$A\hat{\boldsymbol{x}} = proj_{Col\,A}\boldsymbol{b} = \hat{\boldsymbol{b}}$



$$A\widehat{\boldsymbol{x}} = \widehat{\boldsymbol{b}} = proj_{Col\ A}\boldsymbol{b}$$

$$\Updownarrow$$

$$(\boldsymbol{b} - A\widehat{\boldsymbol{x}}) \perp Col\ A = Span\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n\}$$

$$\Updownarrow$$

$$\boldsymbol{a}_j \cdot (\boldsymbol{b} - A\widehat{\boldsymbol{x}}) = \boldsymbol{a}_j^T(\boldsymbol{b} - A\widehat{\boldsymbol{x}}) = 0$$ where $\boldsymbol{a}_j$ is any column of $A = [\boldsymbol{a}_1 \cdots \boldsymbol{a}_n]$

$$\Updownarrow$$

$$A^T(\boldsymbol{b} - A\widehat{\boldsymbol{x}}) = 0$$

$$\Updownarrow$$

$$A^TA\widehat{\boldsymbol{x}} = A^T\boldsymbol{b}$$

Theorem 6.13: The Normal Equations

The set of least-squares solutions of $A\boldsymbol{x} = \boldsymbol{b}$ coincides with the nonempty set of solutions of the normal equations $A^T A\boldsymbol{x} = A^T\boldsymbol{b}$

Diskussion:

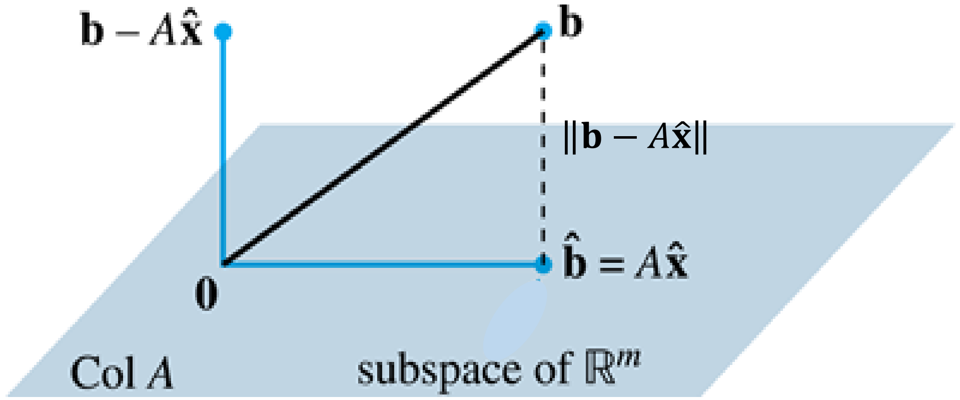
Givet A er en m*n matrix – hvilken dimension har så A^TA?
Hvad skal der mon til for at finde x?

$$\boldsymbol{x} = (A^T A)^{-1} A^T \boldsymbol{b}$$

Matricen foran **b** kaldes:
More-Penrose pseudo-inverse

Ex 4

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \quad \boldsymbol{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$



Dette er et *overbestemt* ligningssystem - kendetegnet ved <u>flere ligninger end ubekendte</u>

Ved løsning af normalligningerne finder vi den *bedst mulige* løsning: $\widehat{\boldsymbol{x}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

Least-squares error: $\|\boldsymbol{b} - A\widehat{\boldsymbol{x}}\|$

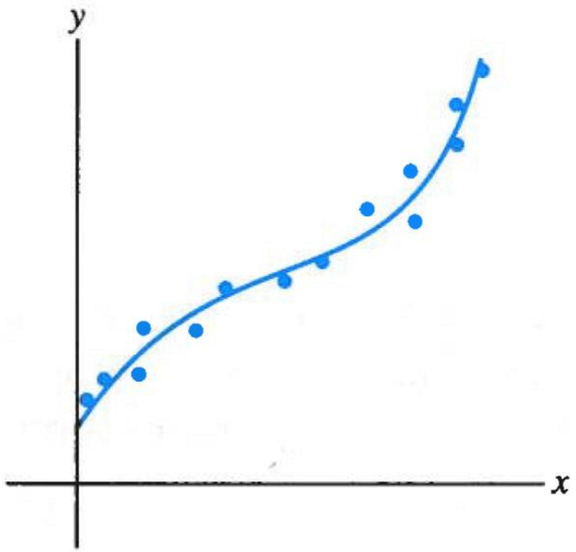
OPGAVE 1

For det overbestemte ligningssystem $A\boldsymbol{x} = \boldsymbol{b}$, hvor:

$$A = \begin{bmatrix} 3 & -2 \\ -3 & 1 \\ -1 & 4 \end{bmatrix} \qquad \boldsymbol{b} = \begin{bmatrix} 6 \\ 1 \\ 9 \end{bmatrix}$$

- Bestem den pseudo-inverse matrix
- Find det **x**, som i mindste kvadraters forstand bedst muligt approximere **b**, når multipliceret med A
- Udregn LSE (least squares error)

6.6 Applications to Linear Models



Least-Squares Lines

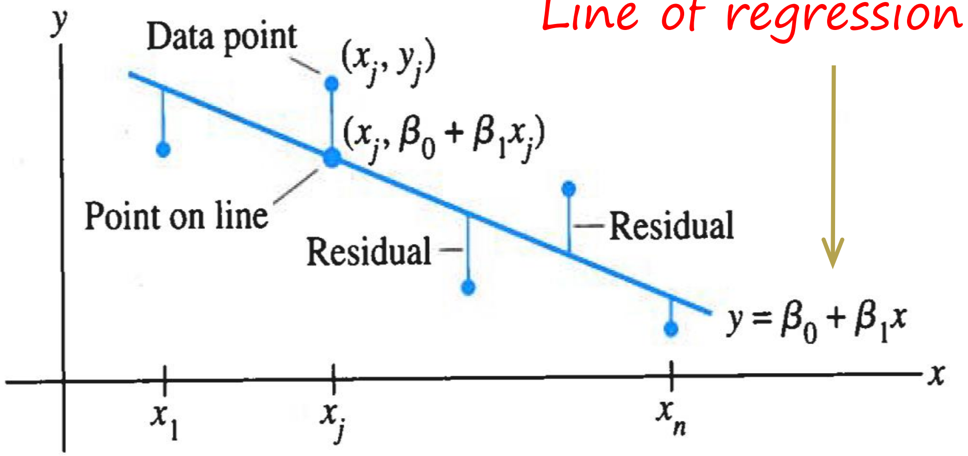
Fitting data to a straight line:

$$y = \beta_0 + \beta_1 x$$

$\hat{y} \rightarrow$

Predicted y-value		Observed y-value
$\beta_0 + \beta_1 x_1$	$=$	y_1
$\beta_0 + \beta_1 x_2$	$=$	y_2
$\vdots$		$\vdots$
$\beta_0 + \beta_1 x_n$	$=$	y_n

Line of regression



Data points: $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i = \hat{y}_i + \varepsilon_i$, where $\hat{y}_i = \beta_0 + \beta_1 x_i$ is the predicted y-value

Residual (error): $\varepsilon_i = y_i - (\beta_0 + \beta_1 x_i) = y_i - \hat{y}_i$

Determine the unknown parameters β_0 and β_1 that minimize the sum of square residuals:

$$\sum_{i=1}^{n} {\varepsilon_i}^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \|\boldsymbol{y} - \hat{\boldsymbol{y}}\|^2$$

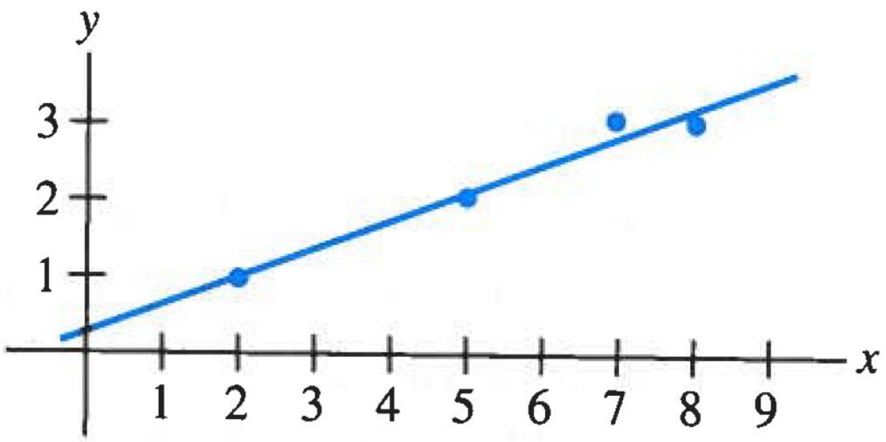
That is, find a least-squares solution of $X\boldsymbol{\beta} = \boldsymbol{y}$, where $X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$, $\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$ and $\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$

Ex 5

Data points: $(x_i, y_i) = (2,1), (5,2), (7,3), (8,3)$ → $y = \beta_0 + \beta_1 x$

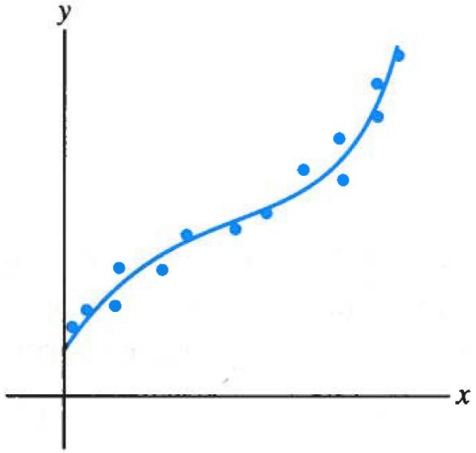
Beregning:

Bestem de koefficienter β_0 og β_1, som for en ret linje giver den bedste approximation til datasættet.



$\beta_0 = 0.286$
$\beta_1 = 0.357$

Least-Squares fitting



Fitting data to polynomium: $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$

Data points: $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$

Determine the unknown parameters $\beta_0, \cdots, \beta_3$ that minimize the sum of square residuals:

$$\sum_{i=1}^{n} {\varepsilon_i}^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \|\boldsymbol{y} - \hat{\boldsymbol{y}}\|^2, \qquad \hat{\boldsymbol{y}} = X\boldsymbol{\beta}$$

That is, find a least-squares solution of: $X\boldsymbol{\beta} = \boldsymbol{y}$

Observation vector

Design matrix

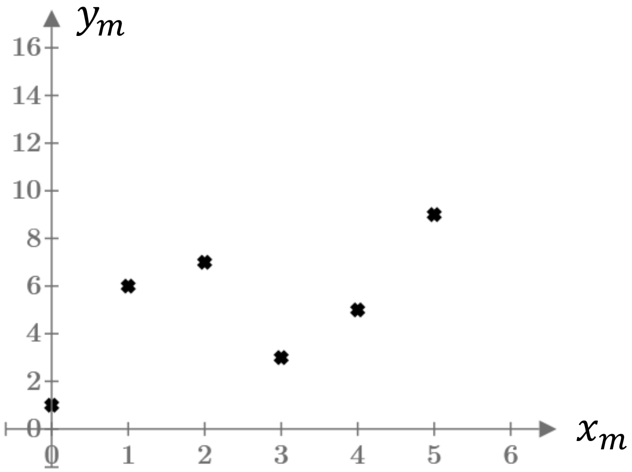
Parameter vector

Residual vector

where: $\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ $\quad X = \begin{bmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & x_n^3 \end{bmatrix}$ $\quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$ $\quad \boldsymbol{\varepsilon} = \boldsymbol{y} - \hat{\boldsymbol{y}} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$

Ex 6

Data points: $x_m := \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$ $\quad y_m := \begin{bmatrix} 1 \\ 6 \\ 7 \\ 3 \\ 5 \\ 9 \end{bmatrix}$



Fitting curve: $f(x) = \beta_0 + \beta_1 \cdot x + \beta_2 \cdot x^2 + \beta_3 \cdot x^3$

Parameter vector: $\beta := \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$

Design matrix: $X := \begin{bmatrix} x_m^{\,0} & x_m^{\,1} & x_m^{\,2} & x_m^{\,3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \\ 1 & 5 & 25 & 125 \end{bmatrix}$

$\Downarrow$

$X^{\mathrm{T}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 4 & 9 & 16 & 25 \\ 0 & 1 & 8 & 27 & 64 & 125 \end{bmatrix} \quad \Longrightarrow \quad X^{\mathrm{T}} X = \begin{bmatrix} 6 & 15 & 55 & 225 \\ 15 & 55 & 225 & 979 \\ 55 & 225 & 979 & 4425 \\ 225 & 979 & 4425 & 20515 \end{bmatrix}$

Ex 6

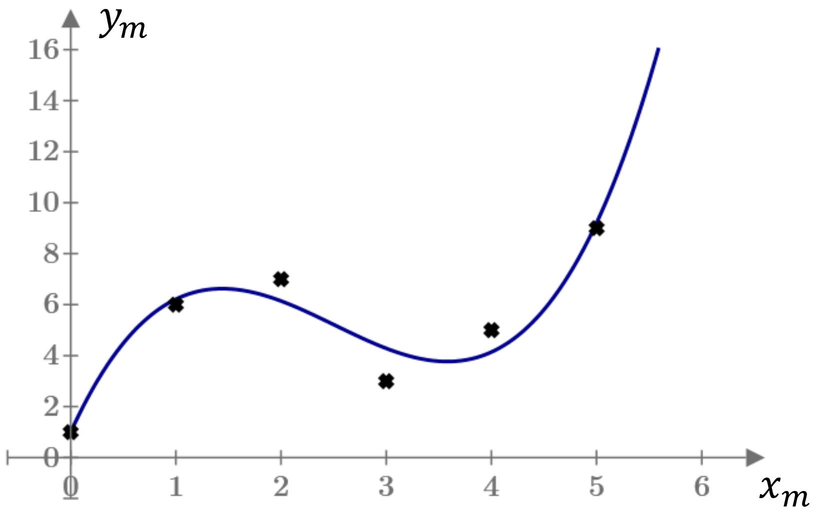
Normal equation: $X^{\mathrm{T}} X \cdot \beta = X^{\mathrm{T}} y_m \implies \begin{bmatrix} 6 & 15 & 55 & 225 \\ 15 & 55 & 225 & 979 \\ 55 & 225 & 979 & 4425 \\ 225 & 979 & 4425 & 20515 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 4 & 9 & 16 & 25 \\ 0 & 1 & 8 & 27 & 64 & 125 \end{bmatrix} \begin{bmatrix} 1 \\ 6 \\ 7 \\ 3 \\ 5 \\ 9 \end{bmatrix} = \begin{bmatrix} 31 \\ 94 \\ 366 \\ 1588 \end{bmatrix}$

Augmented matrix: $A_{aug} = \left[X^{\mathrm{T}} X;\ X^{\mathrm{T}} y_m \right] = \begin{bmatrix} 6 & 15 & 55 & 225 & 31 \\ 15 & 55 & 225 & 979 & 94 \\ 55 & 225 & 979 & 4425 & 366 \\ 225 & 979 & 4425 & 20515 & 1588 \end{bmatrix}$

Reduced echelon form: $A_{rref} := \mathrm{rref}\left(A_{aug}\right) = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 9.024 \\ 0 & 0 & 1 & 0 & -4.393 \\ 0 & 0 & 0 & 1 & 0.583 \end{bmatrix} \implies \beta := \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 9.024 \\ -4.393 \\ 0.583 \end{bmatrix}$

Best regression curve: $f(x) := 1 + 9.024 \cdot x - 4.393 \cdot x^2 + 0.583 \cdot x^3$

Residuals (errors): $\varepsilon := y_m - f\left(x_m\right) = \begin{bmatrix} 0 \\ -0.214 \\ 0.857 \\ -1.286 \\ 0.857 \\ -0.214 \end{bmatrix}$ $\quad \|\varepsilon\| = 1.793$



Vi har som eksempel et datasæt (x,y), som vi ønsker approximeret med en parabolisk funktion

$$y = ax^2 + bx + c$$

Det første datapar (x,y) er (2,3)

Diskussion:

Hvilke talværdier består første række af designmatricen af?

Første række i designmatricen må være [1 2 4]

Least-Squares fitting of other curves

Known (fitting) functions

Fitting data to known functions: $y = \beta_0 f_0(x) + \beta_1 f_1(x) + \cdots + \beta_k f_k(x)$ ← Linear model = Linear in the unknown paratemers β_i

Data points: $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$

Determine the unknown parameters $\beta_0, \cdots, \beta_k$ that minimize the sum of square residuals:

$$\sum_{i=1}^{n} {\varepsilon_i}^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \|\boldsymbol{y} - \hat{\boldsymbol{y}}\|^2, \qquad \hat{\boldsymbol{y}} = X\boldsymbol{\beta}$$

That is, find a least-squares solution of: $X\boldsymbol{\beta} = \boldsymbol{y}$

Observation vector | Design matrix | Parameter vector | Residual vector

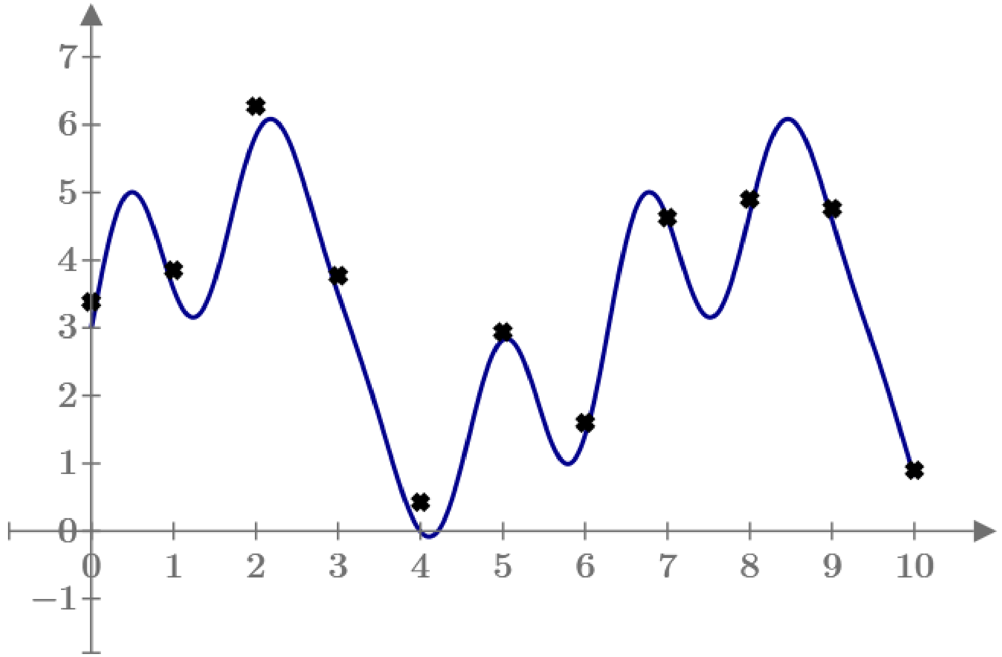
where: $\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ $\quad X = \begin{bmatrix} f_0(x_1) & f_1(x_1) & \cdots & f_k(x_1) \\ f_0(x_2) & f_1(x_2) & \cdots & f_k(x_2) \\ \vdots & \vdots & \vdots & \vdots \\ f_0(x_n) & f_1(x_n) & \cdots & f_k(x_n) \end{bmatrix}$ $\quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}$ $\quad \boldsymbol{\varepsilon} = \boldsymbol{y} - \hat{\boldsymbol{y}} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$

Ex 7 Least-Squares fitting of sinosoidal harmonics

Fitting functions: $f_0(x) = 1$; $f_1(x) = \sin(x)$; $f_2(x) = \sin(2x)$; … ; $f_k(x) = \sin(kx)$

Fitting curve: $y = \beta_0 + \beta_1 \sin(x) + \beta_2 \sin(2x) + \cdots + \beta_k \sin(kx)$

Harmonic/Fourier series



Todays words and concepts

Least-squares solution

Gram-Schmidt proces

The normal equation

Least-squares error

Residual

Least-squares fitting

Observation vector

Design matrix

Parameter vector

Residual vector

Least-squares lines

QR factorization