

Anvendt Lineær Algebra

Kursusgang 3

Anita Abildgaard Sillasen

Institut for Matematiske Fag

Vi betragter et lineært ligningssystem (af m ligninger med n ubekendte)

$$A\mathbf{x} = \mathbf{b}.$$

Ligningssystemet antages at være inkonsistent (ingen løsninger) fordi tallene er fremkommet som måleresultater med målefejl.

Vi skal derfor finde en mindste kvadraters løsning $\hat{\mathbf{x}}$ som opfylder

$$A\hat{\mathbf{x}} = \hat{\mathbf{b}},$$

hvor $\hat{\mathbf{b}}$ er den vektor i $\text{Col } A$ hvis afstand til \mathbf{b} er mindst.

$\hat{\mathbf{b}}$ er altså ortogonal projektionen af \mathbf{b} på $\text{Col } A$.

Vi beregner ikke $\hat{\mathbf{b}}$.

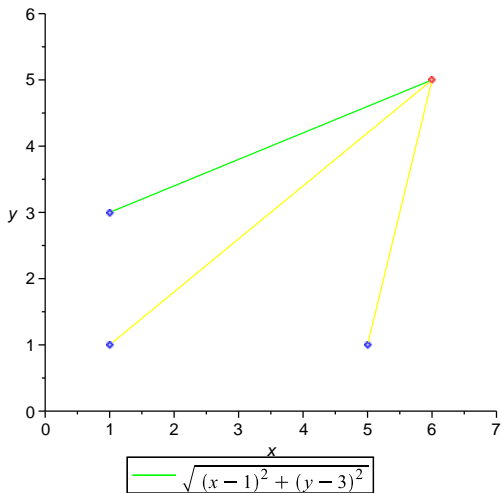
En vektor \mathbf{x} er mindste kvadraters løsning til $A\mathbf{x} = \mathbf{b}$ hvis og kun hvis den er løsning til *normalligningen*

$$(A^T A)\mathbf{x} = A^T \mathbf{b}.$$

Når en mindste kvadraters løsning $\hat{\mathbf{x}}$ er fundet ved at løse normalligningen kan eventuelt bestemme $\hat{\mathbf{b}} = A\hat{\mathbf{x}}$.

Fremgangsmåden virker kun for *lineære* ligninger.

Ikke-lineære problemer.



Vi kender de præcise koordinater for et antal punkter (blå på figuren på forrige side). Desuden er der et ukendt punkt (rødt).

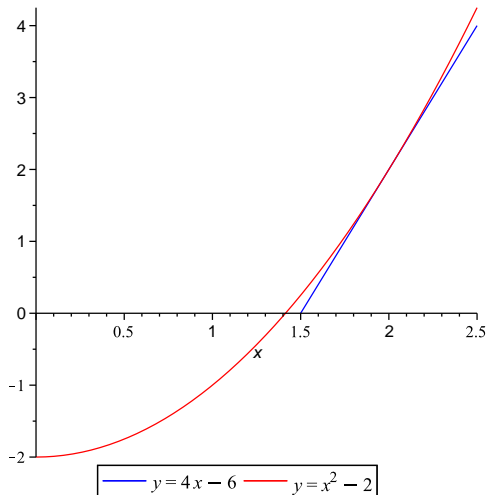
Afstanden mellem det kendte punkt $(1, 3)$ og det ukendte punkt (x, y) måles til værdien b . Vi har altså følgende ligning (hvis vi ser bort fra målefejl):

$$\sqrt{(x - 1)^2 + (y - 3)^2} = b.$$

Tilsvarende ligninger fås ved måling af afstanden til andre kendte punkter.

Men disse ligninger er ikke-lineære.

Lineær Approximation



Lineær approximation af funktion af én variabel

Ligningen for tangenten til grafen for $f(x)$ i punktet $(a, f(a))$ er

$$y = f(a) + f'(a)(x - a).$$

Lineær approximation for x tæt på a :

$$f(x) \approx f(a) + f'(a)(x - a).$$

F.eks. hvis $f(x) = x^2 - 2$, $a = 2$ så er $f'(x) = 2x$
og tangentens ligning er: $y = 2 + 4(x - 2)$ eller $y = 4x - 6$.

Løsning af ligning ved hjælp af lineær approximation (Newtons metode):

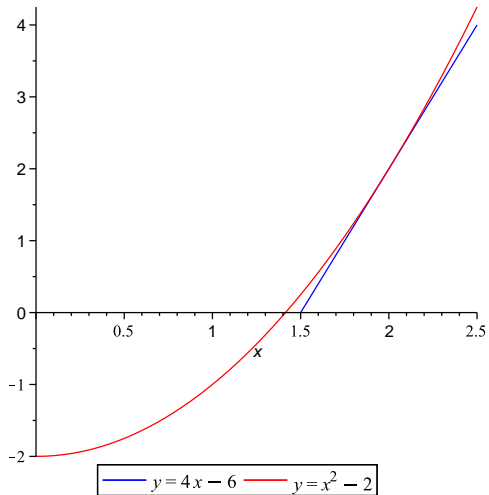
Eksempel. Løs ligning $f(x) = 0$ hvor $f(x) = x^2 - 2$.

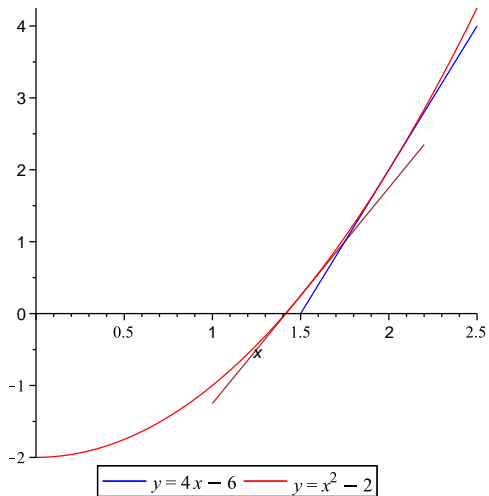
Vi vil bestemme en følge af tal $x^0, x^1, x^2, x^3, \dots$, der konvergerer mod den præcise løsning.

x^0 er vores første gæt på en løsning: $x^0 = 2$.

I nærheden af x^0 er $f(x) \approx f(x^0) + f'(x^0)(x - x^0) = 4x - 6$.

Da løsningen formodes at ligge tæt på x^0 kan vi løse ligningen $4x - 6 = 0$ i stedet for den ikke-lineære ligning. Denne ligning har løsning $x^1 = 1.5$.





I nærheden af x^1 er $f(x) \approx f(x^1) + f'(x^1)(x - x^1) = 3x - 4.25$.

I stedet for den ikke-lineære ligning løser vi ligningen $3x - 4.25 = 0$ som har løsning $x^2 = 1.4166666$.

Dernæst fås

$$x^3 = 1.4142156$$

og

$$x^4 = 1.4142135.$$

Funktion af to variable

$f(x, y)$: en funktion af to variable med partielle afledede

$$f_x(x, y) = \frac{\partial f}{\partial x} \quad \text{og} \quad f_y(x, y) = \frac{\partial f}{\partial y}.$$

Tilvæksten af $f(x, y)$ ud fra punktet (a, b) :

$$\Delta f = f(a + \Delta x, b + \Delta y) - f(a, b)$$

approximeres af *differentialet*

$$df = f_x(a, b)\Delta x + f_y(a, b)\Delta y.$$

Differentialet kan fremkomme som prikprodukt af gradient vektoren

$\nabla f(a, b) = \langle f_x(a, b), f_y(a, b) \rangle$ og vektoren $\langle \Delta x, \Delta y \rangle$:

$$df = \nabla f(a, b) \cdot \langle \Delta x, \Delta y \rangle.$$

Lineær approximation for funktion af to variable

$$f(a + \Delta x, b + \Delta y) \approx f(a, b) + f_x(a, b)\Delta x + f_y(a, b)\Delta y,$$

Eller, når (x, y) er tilstrækkeligt tæt på (a, b) :

$$f(x, y) \approx f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

Lad $f(x, y)$ være afstanden mellem (x, y) og et fast punkt (x_1, y_1) .

$$f(x, y) = \sqrt{(x - x_1)^2 + (y - y_1)^2}.$$

Så er

$$f_x(x, y) = \frac{x - x_1}{\sqrt{(x - x_1)^2 + (y - y_1)^2}}, \quad f_y(x, y) = \frac{y - y_1}{\sqrt{(x - x_1)^2 + (y - y_1)^2}}.$$

For (x, y) i nærheden af punktet (x^0, y^0) er

$$f(x, y) \approx f(x^0, y^0) + f_x(x^0, y^0)(x - x^0) + f_y(x^0, y^0)(y - y^0).$$

Vi kender de præcise koordinater for et antal punkter: (x_1, y_1) , (x_2, y_2) ,
 $\dots, (x_n, y_n)$.

Desuden er der et punkt med ukendte koordinater (x_P, y_P) .

Vi har målt afstanden fra dette punkt til de n kendte punkter og får værdierne henholdsvis b_1, b_2, \dots, b_n .

Altså:

$$\begin{aligned}\sqrt{(x_P - x_1)^2 + (y_P - y_1)^2} &= b_1 \\ \sqrt{(x_P - x_2)^2 + (y_P - y_2)^2} &= b_2 \\ &\vdots \\ \sqrt{(x_P - x_n)^2 + (y_P - y_n)^2} &= b_n\end{aligned}$$

P.g.a. målefejl er der ikke noget punkt (x_P, y_P) der passer med alle ligninger.

Vi ønsker at finde et punkt, der passer bedst muligt.

Idé:

Gæt et punkt (x_P^0, y_P^0) .

Da (x_P, y_P) formodes at være tæt på (x_P^0, y_P^0) kan venstre side i ovenstående ligninger erstattes med den lineære approximation i punktet (x_P^0, y_P^0) .

Derved fås (inkonsistent) lineært ligningssystem med n ligninger. Lad (x_P^1, y_P^1) være en mindste kvadraters løsning til dette ligningssystem.

(x_P^1, y_P^1) er så forhåbentligt tættere på den rigtige løsning end vores første gæt var.

Gentag ovenstående med (x_P^0, y_P^0) erstattet af (x_P^1, y_P^1) og find derved (x_P^2, y_P^2) .

Gentag indtil (x_P^i, y_P^i) nærmer sig et bestemt punkt, som så er (x_P, y_P) .

Funktion af n variable

$F_i(x_1, x_2, \dots, x_n)$: en funktion af n variable.

$\mathbf{x}^0 = (x_1^0, x_2^0, \dots, x_n^0)$: fast punkt/vektor.

$\mathbf{j}_i = (j_{i1}, j_{i2}, \dots, j_{in}) = \nabla F_i(\mathbf{x}^0)$ er gradientvektoren af F_i i punktet \mathbf{x}^0 , hvor $j_{i\ell}$ er $\frac{\partial F_i}{\partial x_\ell}$ udregnet i punktet \mathbf{x}^0 .

For et punkt $\mathbf{x} = (x_1, x_2, \dots, x_n)$ i nærheden af \mathbf{x}^0 er

$$F_i(\mathbf{x}) \approx F_i(\mathbf{x}^0) + \nabla F_i(\mathbf{x}^0) \cdot (\mathbf{x} - \mathbf{x}^0) = b_i^0 + \mathbf{j}_i \cdot (\mathbf{x} - \mathbf{x}^0),$$

hvor $b_i^0 = F_i(\mathbf{x}^0)$.

Vi har m observationer:

$$F_1(x_1, \dots, x_n) = b_1$$

$$F_2(x_1, \dots, x_n) = b_2$$

\vdots

$$F_m(x_1, \dots, x_n) = b_m.$$

Venstresiderne erstattes af deres lineære approximationer i punktet \mathbf{x}^0 :

$$b_1^0 + \mathbf{j}_1 \cdot (\mathbf{x} - \mathbf{x}^0) = b_1$$

$$b_2^0 + \mathbf{j}_2 \cdot (\mathbf{x} - \mathbf{x}^0) = b_2$$

\vdots

$$b_m^0 + \mathbf{j}_m \cdot (\mathbf{x} - \mathbf{x}^0) = b_m.$$

$$\mathbf{j}_1 \cdot (\mathbf{x} - \mathbf{x}^0) = b_1 - b_1^0$$

$$\vdots$$

$$\mathbf{j}_m \cdot (\mathbf{x} - \mathbf{x}^0) = b_m - b_m^0.$$

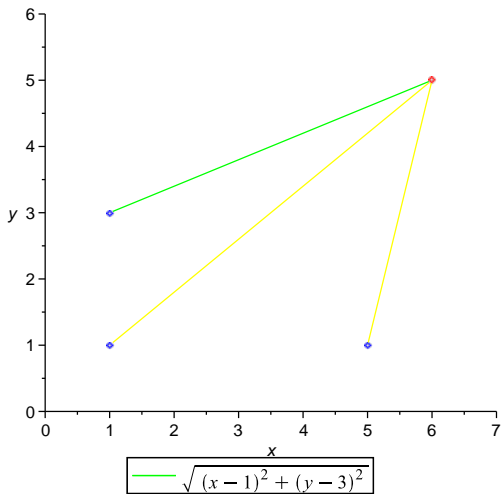
På matrixform: $A(\mathbf{x} - \mathbf{x}^0) = \mathbf{b}$, hvor

$$A = \begin{bmatrix} j_{11} & j_{12} & \cdots & j_{1n} \\ j_{21} & j_{22} & \cdots & j_{2n} \\ \vdots & \vdots & & \vdots \\ j_{m1} & j_{m2} & \cdots & j_{mn} \end{bmatrix} \quad \text{og} \quad \mathbf{b} = \begin{bmatrix} b_1 - b_1^0 \\ b_2 - b_2^0 \\ \vdots \\ b_m - b_m^0 \end{bmatrix}.$$

Med residualvektoren $\hat{\mathbf{r}}$ fås observationsligningen

$$A(\mathbf{x} - \mathbf{x}^0) = \mathbf{b} + \hat{\mathbf{r}}.$$

A kaldes designmatricen.



Der er tre kendte punkter: $(1, 3)$, $(1, 1)$ og $(5, 1)$
samt et ukendt punkt (x_P, y_P) med afstande til de tre kendte punkter målt til
henholdsvis 4, 6 og 5.

$$F_1(x, y) := \sqrt{(x - 1)^2 + (y - 3)^2}$$

$$F_2(x, y) := \sqrt{(x - 1)^2 + (y - 1)^2}$$

$$F_3(x, y) := \sqrt{(x - 5)^2 + (y - 1)^2}$$

$$F_1(x_P, y_P) = 4, \quad F_2(x_P, y_P) = 6, \quad F_3(x_P, y_P) = 5.$$

$$\mathbf{j}_1 = \nabla F_1(x_P^0, y_P^0) = \left(\frac{x_P^0 - 1}{\sqrt{(x_P^0 - 1)^2 + (y_P^0 - 3)^2}}, \frac{y_P^0 - 3}{\sqrt{(x_P^0 - 1)^2 + (y_P^0 - 3)^2}} \right) = \left(\frac{x_P^0 - 1}{b_1^0}, \frac{y_P^0 - 3}{b_1^0} \right)$$

$$\mathbf{j}_2 = \nabla F_2(x_P^0, y_P^0) = \left(\frac{x_P^0 - 1}{b_2^0}, \frac{y_P^0 - 1}{b_2^0} \right)$$

$$\mathbf{j}_3 = \nabla F_3(x_P^0, y_P^0) = \left(\frac{x_P^0 - 5}{b_3^0}, \frac{y_P^0 - 1}{b_3^0} \right),$$

hvor $b_i^0 = F_i(x_P^0, y_P^0)$.

Vi gætter en første værdi $(x_P^0, y_P^0) = (6, 5)$ og udregner

$$b_1^0 = F_1(6, 5) = \sqrt{(6-1)^2 + (5-3)^2} = \sqrt{29}$$

$$b_2^0 = F_2(6, 5) = \sqrt{(6-1)^2 + (5-1)^2} = \sqrt{41}$$

$$b_3^0 = F_1(6, 5) = \sqrt{(6-5)^2 + (5-1)^2} = \sqrt{17}$$

$$\mathbf{j}_1 = \left(\frac{6-1}{\sqrt{29}}, \frac{5-3}{\sqrt{29}} \right) = (0.927, 0.372)$$

$$\mathbf{j}_2 = \left(\frac{6-1}{\sqrt{41}}, \frac{5-1}{\sqrt{41}} \right) = (0.781, 0.625)$$

$$\mathbf{j}_3 = \left(\frac{6-5}{\sqrt{17}}, \frac{5-1}{\sqrt{17}} \right) = (0.242, 0.968)$$

$$\mathbf{b} = \begin{bmatrix} 4 \\ 6 \\ 5 \end{bmatrix} - \begin{bmatrix} \sqrt{29} \\ \sqrt{41} \\ \sqrt{17} \end{bmatrix} = \begin{bmatrix} -1.39 \\ -0.40 \\ 0.88 \end{bmatrix}$$

Observationsligning (uden residualer):

$$\begin{bmatrix} 0.927 & 0.372 \\ 0.781 & 0.625 \\ 0.242 & 0.968 \end{bmatrix} (\mathbf{x} - \mathbf{x}_0) = \begin{bmatrix} -1.39 \\ -0.40 \\ 0.88 \end{bmatrix}.$$

Normal ligning ($A^T A(\mathbf{x} - \mathbf{x}_0) = A^T \mathbf{b}$):

$$\begin{bmatrix} 1.531 & 1.068 \\ 1.068 & 1.469 \end{bmatrix} (\mathbf{x} - \mathbf{x}_0) = \begin{bmatrix} -1.388 \\ 0.0847 \end{bmatrix}.$$

Løsning:

$$\mathbf{x} - \mathbf{x}_0 = \begin{bmatrix} -1.92 \\ 1.46 \end{bmatrix}.$$

Altså:

$$\mathbf{x}_1 = \begin{bmatrix} x_P^1 \\ y_P^1 \end{bmatrix} = \begin{bmatrix} x_P^0 \\ y_P^0 \end{bmatrix} + \begin{bmatrix} -1.92 \\ 1.46 \end{bmatrix} = \begin{bmatrix} 4.078 \\ 6.454 \end{bmatrix}.$$

Gentag med (x_P^0, y_P^0) erstattet af (x_P^1, y_P^1) .

$$\text{Så fås: } \begin{bmatrix} x_P^2 \\ y_P^2 \end{bmatrix} = \begin{bmatrix} x_P^1 \\ y_P^1 \end{bmatrix} + \begin{bmatrix} -0.113 \\ -0.496 \end{bmatrix} = \begin{bmatrix} 3.965 \\ 5.958 \end{bmatrix}.$$

$$\text{Dernæst: } \begin{bmatrix} x_P^3 \\ y_P^3 \end{bmatrix} = \begin{bmatrix} x_P^2 \\ y_P^2 \end{bmatrix} + \begin{bmatrix} -0.007 \\ -0.00003 \end{bmatrix} = \begin{bmatrix} 3.958 \\ 5.958 \end{bmatrix}.$$

Så er

$$F_1(x_P^3, y_P^3) = 4.183, F_2(x_P^3, y_P^3) = 5.773, F_3(x_P^3, y_P^3) = 5.066.$$

Vægtet mindste kvadraters metode

For et lineært ligningssystem (af m ligninger med n ubekendte)

$$A\mathbf{x} = \mathbf{b}$$

er en mindste kvadraters løsning $\hat{\mathbf{x}}$ en løsning til *normalligningen*

$$(A^T A)\mathbf{x} = A^T \mathbf{b}.$$

Hvis

$$A = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \quad \text{og} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

så er $A^T A = m$ og $A^T \mathbf{b} = b_1 + \dots + b_m$ og derfor er

$$\hat{x} = \frac{b_1 + \dots + b_m}{m},$$

altså (ikke-vægtet) gennemsnit af b_1, \dots, b_m .

Det vægtede gennemsnit af b_1, \dots, b_m med vægte hhv. c_1, \dots, c_m (hvor c_1, \dots, c_m er positive tal) er

$$\frac{c_1 b_1 + \dots + c_m b_m}{c_1 + \dots + c_m}.$$

Betragt nu et ligningssystem af m ligninger med n ubekendte

$$A\mathbf{x} = \mathbf{b}.$$

De m ligninger (observationer) tildeles vægte c_1, \dots, c_m som skrives i en matrix

$$C = \begin{bmatrix} c_1 & 0 & \dots & 0 \\ 0 & c_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & c_m \end{bmatrix}.$$

Som vægte kan vi bruge $c_i = \frac{1}{\sigma_i^2}$ hvor σ_i^2 er variansen af den i 'te observation, eller eventuelt $c_i = \text{konst.} \cdot \frac{1}{\sigma_i^2}$.

Skrives også (i Lay) som $w_i^2 = c_i$ hvor $w_i = \frac{1}{\sigma_i}$ og

$$W = \begin{bmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & w_m \end{bmatrix}.$$

Så er $W^T W = W W = C$.

Hvis vægtene c_1, \dots, c_m er hele tal (positive) så svarer det vægtede ligningssystem til at betragte (ikke-vægtet) ligningssystem

$$\tilde{\mathbf{A}}\mathbf{x} = \tilde{\mathbf{b}},$$

hvor ligning nr. i fra ligningssystemet $\mathbf{A}\mathbf{x} = \mathbf{b}$ skrives c_i gange.

En mindste kvadraters løsning til det nye ligningssystem er en løsning til normalligningen

$$\tilde{\mathbf{A}}^T \tilde{\mathbf{A}}\mathbf{x} = \tilde{\mathbf{A}}^T \tilde{\mathbf{b}}.$$

Eksempel:

$$A = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} r \\ s \\ t \end{bmatrix}, \quad C = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Så kan vi betragte

$$\tilde{A} = \begin{bmatrix} a & b \\ a & b \\ c & d \\ e & f \\ e & f \\ e & f \end{bmatrix}, \quad \tilde{\mathbf{b}} = \begin{bmatrix} r \\ r \\ s \\ t \\ t \\ t \end{bmatrix}.$$

Normalligningen for ligningssystemet $\tilde{A}\mathbf{x} = \tilde{\mathbf{b}}$ er

$$\tilde{A}^T \tilde{A}\mathbf{x} = \tilde{A}^T \tilde{\mathbf{b}}.$$

Da $\tilde{A}^T \tilde{A} = A^T C A$ og $\tilde{A}^T \tilde{\mathbf{b}} = A^T C \mathbf{b}$ er denne ligning ækvivalent med

$$A^T C A \mathbf{x} = A^T C \mathbf{b}.$$

Det vægtede ligningsystem $A\mathbf{x} = \mathbf{b}$ med generelle vægte har *normalligning*

$$A^T C A \mathbf{x} = A^T C \mathbf{b}.$$

En *vægtet mindste kvadraters løsning* til $A\mathbf{x} = \mathbf{b}$ er en løsning til normalligningen.

I observationsligningen $A\hat{\mathbf{x}} = \mathbf{b} + \hat{\mathbf{r}}$ indgår residual vektoren

$$\hat{\mathbf{r}} = \begin{bmatrix} \hat{r}_1 \\ \vdots \\ \hat{r}_m \end{bmatrix}.$$

Vi udregner

$$\hat{\mathbf{r}}^T C \hat{\mathbf{r}} = [\hat{r}_1 \dots \hat{r}_m] \begin{bmatrix} c_1 & 0 & \dots & 0 \\ 0 & c_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & c_m \end{bmatrix} \begin{bmatrix} \hat{r}_1 \\ \vdots \\ \hat{r}_m \end{bmatrix} =$$
$$[\hat{r}_1 \dots \hat{r}_m] \begin{bmatrix} c_1 \hat{r}_1 \\ \vdots \\ c_m \hat{r}_m \end{bmatrix} = c_1 \hat{r}_1^2 + \dots + c_m \hat{r}_m^2.$$

Dette er kvadratet på den vægtede længde af vektoren $\hat{\mathbf{r}}$. Altså kvadratet på den vægtede afstand mellem $A\hat{\mathbf{x}}$ og \mathbf{b} .

En vægtet mindste kvadraters løsning er en vektor \mathbf{x} som minimierer

$$\hat{\mathbf{r}}^T C \hat{\mathbf{r}} = (\mathbf{Ax} - \mathbf{b})^T C (\mathbf{Ax} - \mathbf{b}).$$

Eksempel.

Hvis

$$A = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \quad \text{og} \quad C = \begin{bmatrix} c_1 & 0 & \dots & 0 \\ 0 & c_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & c_m \end{bmatrix}$$

så er $A^T C A = c_1 + \dots + c_m$ og $A^T C \mathbf{b} = c_1 b_1 + \dots + c_m b_m$ og den vægtede mindste kvadraters løsning er derfor det vægtede gennemsnit

$$\hat{x} = \frac{c_1 b_1 + \dots + c_m b_m}{c_1 + \dots + c_m}.$$

Eksempel.

Find linie $y = \alpha x + \beta$ der passer bedst til punkterne

$$(1, 3), (2, 5), (3, 9), (4, 12), (5, 14)$$

med vægte hhv. 1, 5, 1, 5, 1.

Vi betragter ligningssystemet

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} \beta \\ \alpha \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 9 \\ 12 \\ 14 \end{bmatrix}.$$

Normalligningen:

$$(A^T CA) \begin{bmatrix} \beta \\ \alpha \end{bmatrix} = A^T C \mathbf{b}.$$

Den vægtede mindste kvadraters løsning er

$$\begin{bmatrix} \beta \\ \alpha \end{bmatrix} = (A^T CA)^{-1} A^T C \mathbf{b} = \begin{bmatrix} -0.96 \\ 3.17 \end{bmatrix}.$$

Den vægtede mindste kvadraters linie har ligning

$$y = 3.17x - 0.96.$$

Denne linie er grøn på næste side.

Den ikke-vægtede mindste kvadraters linie er rød på næste side.

