Lineær algebra
Lektion 4

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 Col A hvis afstand til $\mathbf{b}$ er mindst.

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

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

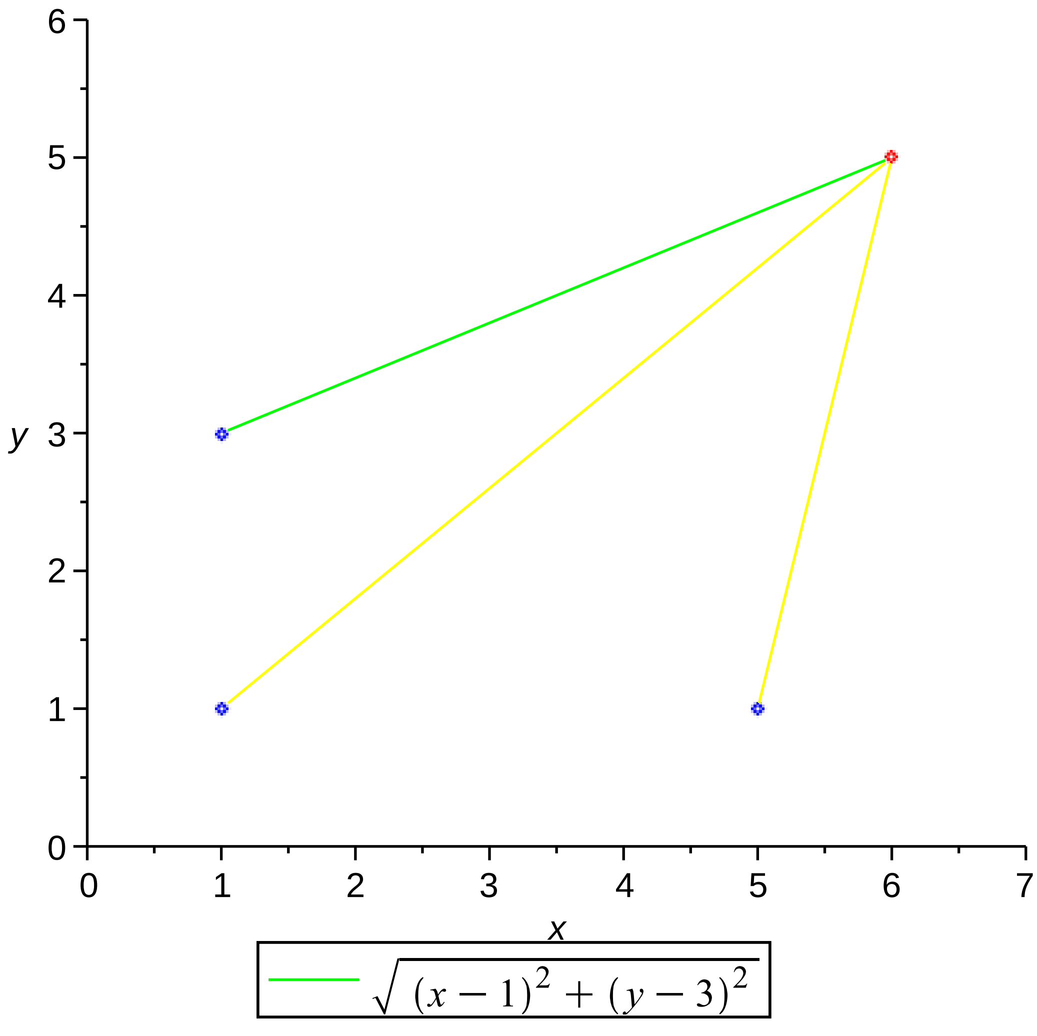
En vektor $\mathbf{x}$ er mindste kvadraters løsning til $A\mathbf{x} = \mathbf{b}$ hvis og kun hvis den er løsning til *normalligningen*

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

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.



6
5
4
y
3
2
1
0
0
1
2
3
4
5
6
7
x
$\sqrt{(x-1)^2+(y-3)^2}$

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

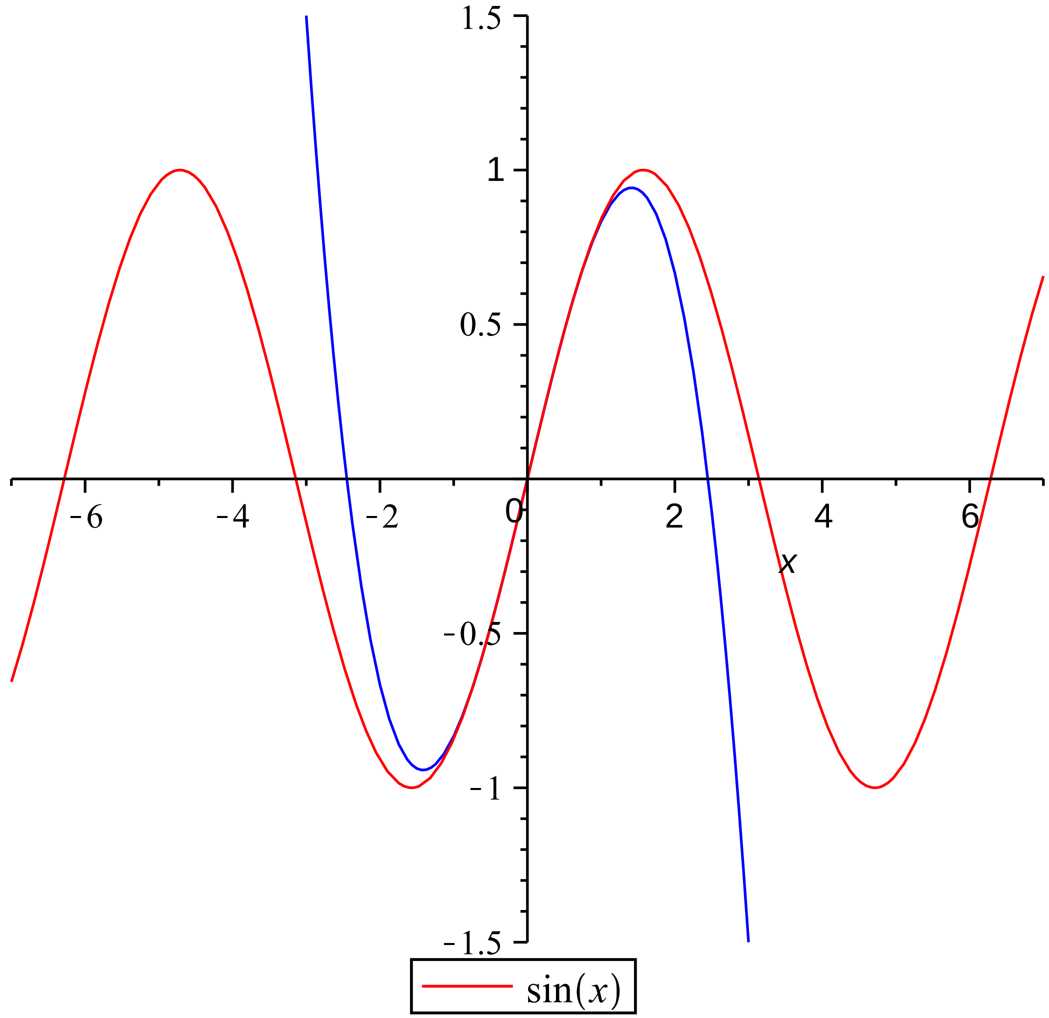
Afstanden mellem det kendte punkt $(1,3)$ og det ukendte punkt (x,y) måles til værdien b. Vi har altså følgende ligning (hvis vi ser bort fra målefejl):

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

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

Men disse ligninger er ikke-lineære.

(Lineær) Approximation


1.5
1
0.5
-0.5
-1
-1.5
-6
-4
-2
0
2
4
6
x
sin(x)

Funktion af én variabel

En tilstrækkeligt pæn funktion $f(x)$ kan i et interval omkring $x = a$ approximeres af Taylorpolynomiet af grad n:

$$f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

F.eks. for $f(x) = \sin x$ og $a = 0$ ses Taylorpolynomiet af grad 3

$$x - \frac{x^3}{6}$$

på figuren på forrige side.

Vi ser kun på lineær approximation – altså Taylorpolynomier af grad 1: $f(a) + f'(a)(x-a)$.

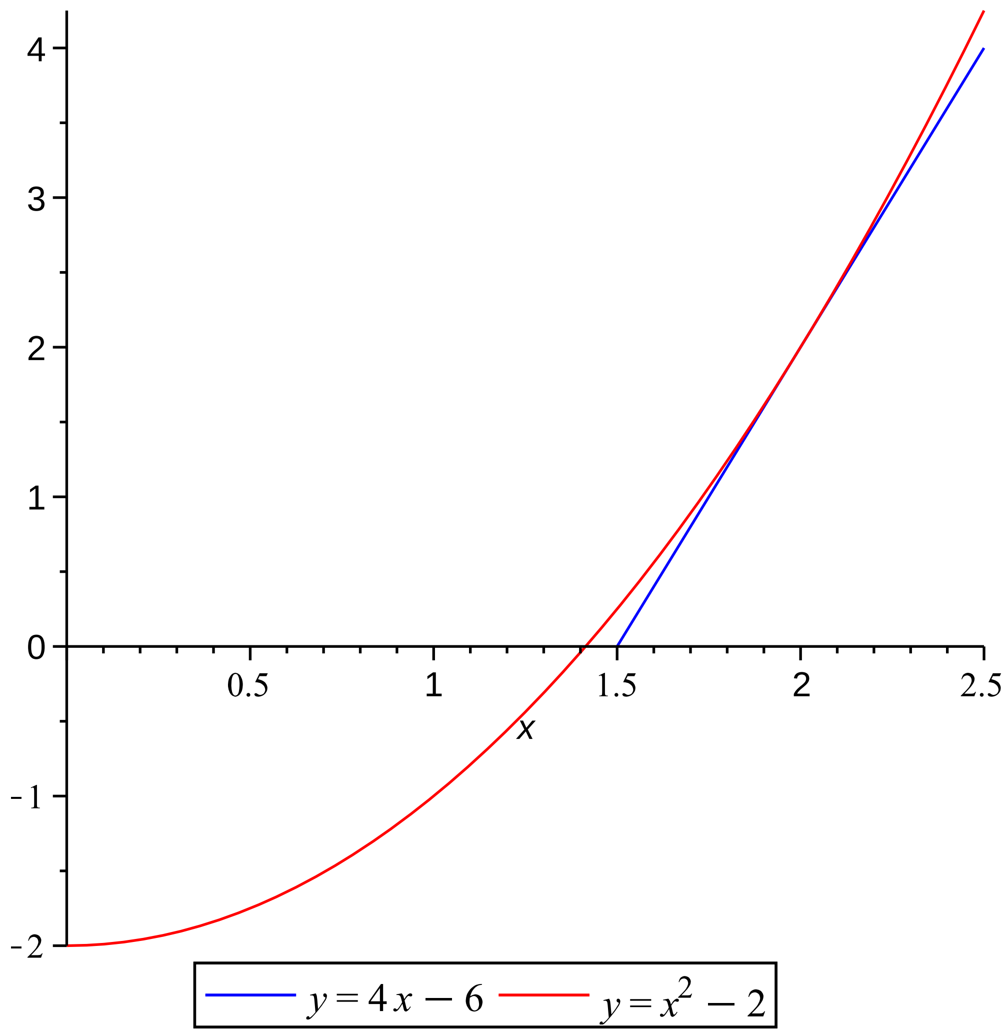
Ligningen for tangenten til grafen for $f(x)$ i punktet $(a, f(a))$ er

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

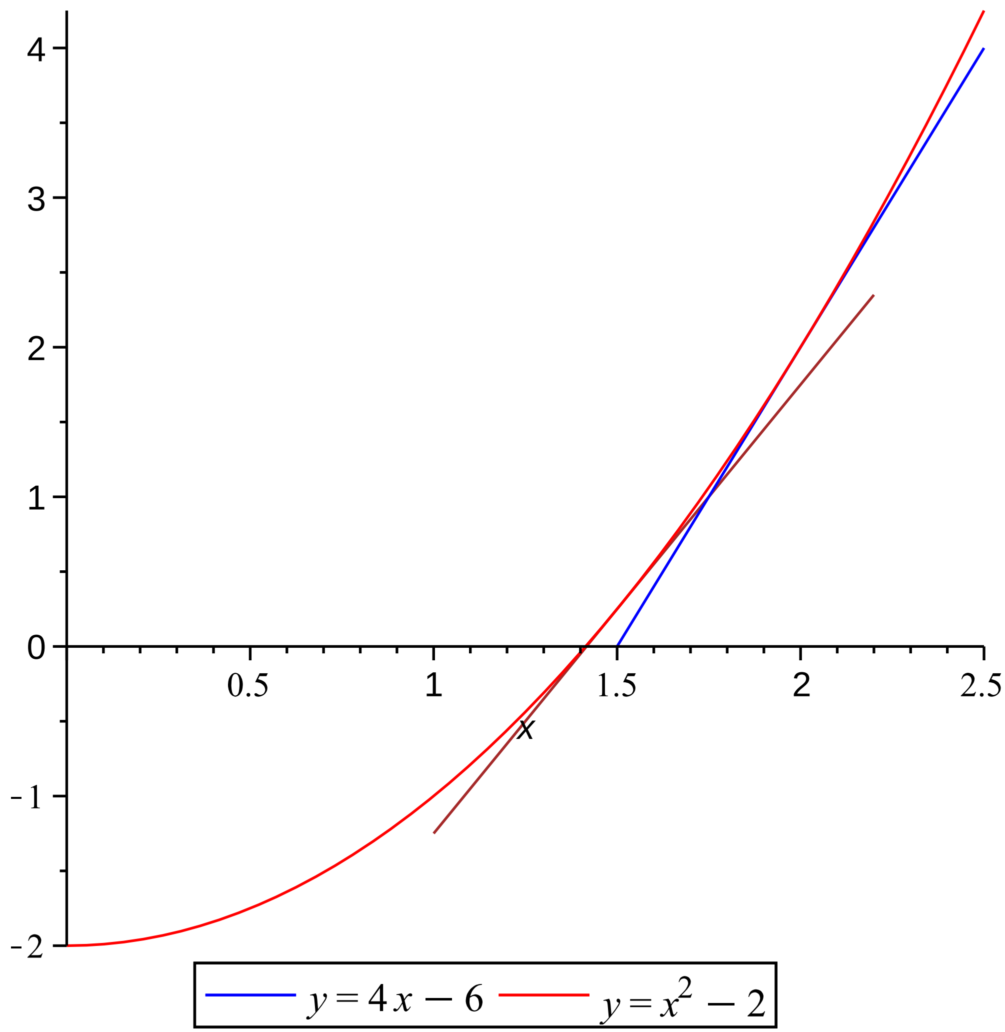
Lineær approximation for x tæt på a:

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

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 $y = 4x - 6$.



4
3
2
1
0
-1
-2
0.5
1
1.5
2
2.5
x
$y = 4x - 6$
$y = x^2 - 2$


4
3
2
1
0
-1
-2
0.5
1
1.5
2
2.5
x
$y = 4x - 6$
$y = x^2 - 2$

Eksempel. Løsning af ikke-lineær ligning $f(x) = 0$ hvor $f(x) = x^2 - 2$. Vi vil bestemme en følge af tal $x^0, x^1, x^2, x^3, \ldots,$ 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

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

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), $\ldots$,(x_n, y_n).

Desuden er der et punkt med ukendte koordinater (x_P, y_P).
Vi har målt afstanden fra dette punkter de n kendte punkter og får værdierne henholdsvis $b_1, b_2, \ldots, b_n$.
Altså:

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

P.g.a. målefejl er der ikke noget punkt (x_P, y_P) der passer med alle ligninger.
Vi ønsker at finder 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 ligningssytem.

(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

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

$\mathbf{x}^0 = (x_1^0, x_2^2, \ldots, x_n^0)$: fast punkt/vektor.

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

For et punkt $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ 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:

$F_1(x_1, \ldots, x_n) = b_1$
$F_2(x_1, \ldots, x_n) = b_2$
$\vdots$
$F_m(x_1, \ldots, x_n) = b_m.$

Venstresiderne erstattes af deres lineære approximationer:

$b_1^0 + \mathbf{j}_1 \cdot (\mathbf{x} - \mathbf{x}^0) = b_1$
$b_2^0 + \mathbf{j}_2 \cdot (\mathbf{x} - \mathbf{x}^0) = b_2$
$\vdots$
$b_m^0 + \mathbf{j}_m \cdot (\mathbf{x} - \mathbf{x}^0) = b_m.$

$$
\begin{aligned}
&\mathbf{j}_1\cdot(\mathbf{x}-\mathbf{x}^0) = b_1 - b_1^0\\
&\vdots\\
&\mathbf{j}_m\cdot(\mathbf{x}-\mathbf{x}^0) = b_m - b_m^0.
\end{aligned}
$$

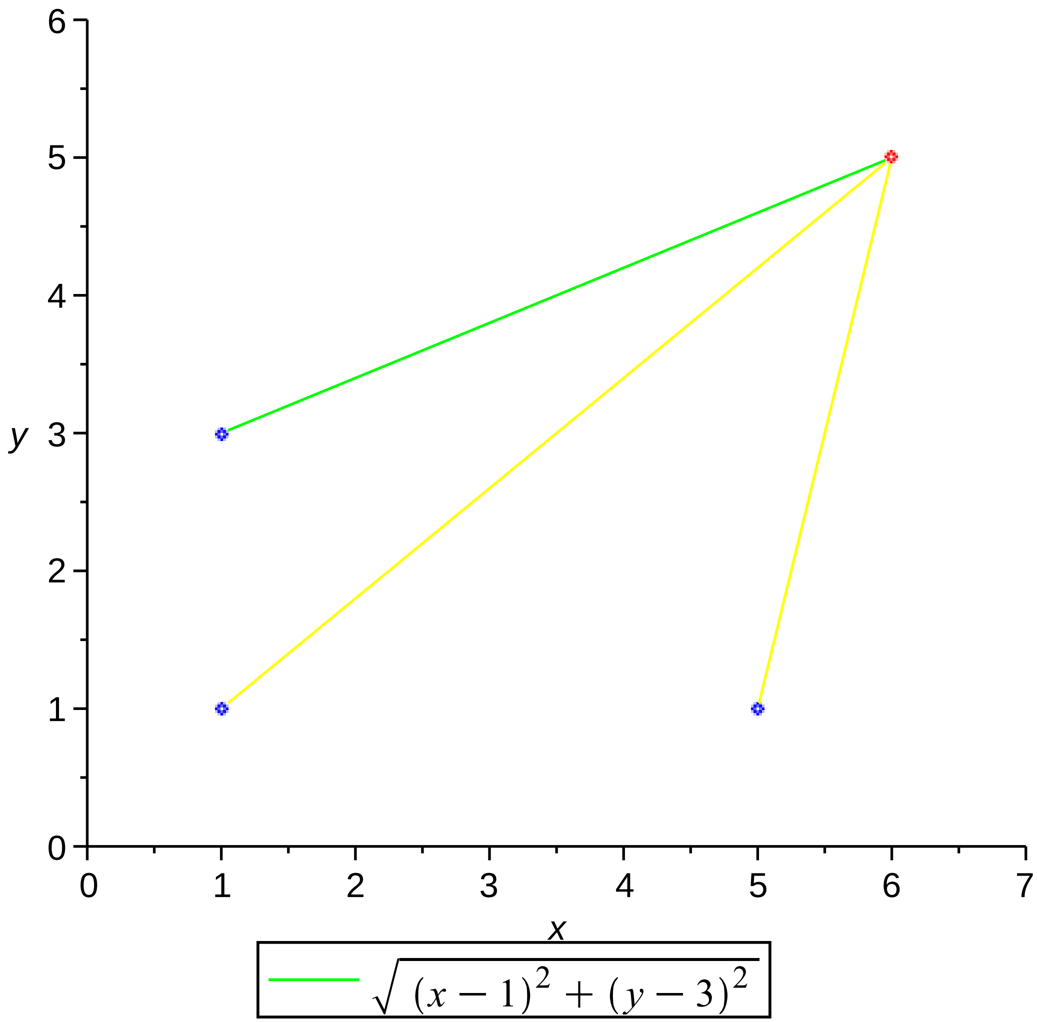
På matrixform: $A(\mathbf{x}-\mathbf{x}^0) = \mathbf{b}$, hvor

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

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

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

A kaldes designmatricen.



6
5
4
y
3
2
1
0
0
1
2
3
4
5
6
7
x
$\sqrt{(x-1)^2+(y-3)^2}$

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

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

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

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

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

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

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

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

Observationsligning (uden residualer):

$$\begin{bmatrix} 0.927 & 0.372 \\ 0.781 & 0.625 \\ 0.242 & 0.968 \end{bmatrix} (\mathbf{x} - \mathbf{x_0}) = \begin{bmatrix} -1.39 \\ -0.40 \\ 0.88 \end{bmatrix}.$$

Normalligning $(A^T A)(\mathbf{x} - \mathbf{x_0}) = A^T \mathbf{b})$:

$$\begin{bmatrix} 1.531 & 1.068 \\ 1.068 & 1.469 \end{bmatrix} (\mathbf{x} - \mathbf{x_0}) = \begin{bmatrix} -1.388 \\ 0.0847 \end{bmatrix}.$$

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).

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

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