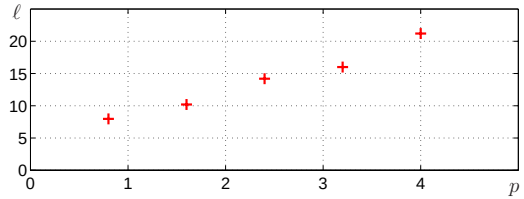
4. Lineære mindste kvadraters problemer pp 55 – 60

Eksempel 4.1. Målt længden ℓ_i for forskllige belastninger p_i af en fjeder.

i	p_i	ℓ_i
1	0.8	7.97
2	1.6	10.2
3	2.4	14.2
4	3.2	16.0
5	4.0	21.2



Punkterne ligger omkring en ret linie $\ell = x_1 + x_2 p$ (Hookes lov)

Ønsker at bestemme parametrene x_1 og x_2

Data fitting: Givet $(t_1, y_1), (t_2, y_2), \ldots, (t_n, y_n)$, som antages at opfylde $y_i = Y(t_i) + \varepsilon_i$, hvor $Y : \mathbb{R} \mapsto \mathbb{R}$ er en vis "baggrunds-funktion" og $\{\varepsilon_i\}$ er "støj", fx målefejl.

Videre har vi givet (eller vælger) en fitte model

$$M(\boldsymbol{x}, t) = x_1 f_1(t) + x_2 f_2(t) + \cdots + x_n f_n(t)$$

hvor $\{f_j\}$ er kendte funktioner og $n \le m$. Ofte $n \ll m$.

Vi ønsker at bestemme parametrene $\{x_j\}$, dvs n-vektoren $\boldsymbol{x}$, så $M(\boldsymbol{x}, t)$ er den "bedst mulige" tilnærmelse til $Y(t)$ i intervallet $\min\{t_i\} \le t \le \max\{t_i\}$

"Bedst mulig" ? Udnyt tilgængelig information.

Residuer
$$\begin{aligned} r_i(\boldsymbol{x}) &= y_i - M(\boldsymbol{x}, t_i) \\ &= y_i - (x_1 f_1(t_i) + x_2 f_2(t_i) + \cdots + x_n f_n(t_i)) \end{aligned}$$

Mindste kvadrater: Bestem $\hat{\boldsymbol{x}}$, det sæt parametre, som minimerer kvadratsummen af residuerne, dvs minimerer funktionen $\varphi(\boldsymbol{x}) = r_1(\boldsymbol{x})^2 + r_2(\boldsymbol{x})^2 + \cdots + r_m(\boldsymbol{x})^2$

Udnytter at parametrene optræder lineært :

$$\boldsymbol{r}(\boldsymbol{x}) = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} - \begin{pmatrix} f_1(t_1)x_1 + f_2(t_1)x_2 + \cdots + f_n(t_1)x_n \\ f_1(t_2)x_1 + f_2(t_2)x_2 + \cdots + f_n(t_2)x_n \\ \vdots \\ f_1(t_m)x_1 + f_2(t_m)x_2 + \cdots + f_n(t_m)x_n \end{pmatrix} = \boldsymbol{y} - \boldsymbol{F}\boldsymbol{x}$$

hvor $\boldsymbol{F}$ er $m \times n$ matricen

$$\boldsymbol{F} = \begin{pmatrix} f_1(t_1) & f_2(t_1) & \cdots & f_n(t_1) \\ f_1(t_2) & f_2(t_2) & \cdots & f_n(t_2) \\ \vdots & & & \\ f_1(t_m) & f_2(t_m) & \cdots & f_n(t_m) \end{pmatrix}$$

Eksempel 4.2. Fjeder problem. Last omdøbes fra p til t. $M(\boldsymbol{x}, t) = x_1 + x_2 t$, dvs $f_1(t) = 1$, $f_2(t) = t$

$$\boldsymbol{y} = \begin{pmatrix} 7.97 \\ 10.2 \\ 14.2 \\ 16.0 \\ 21.2 \end{pmatrix}, \quad \boldsymbol{F} = \begin{pmatrix} 1 & 0.8 \\ 1 & 1.6 \\ 1 & 2.4 \\ 1 & 3.2 \\ 1 & 4.0 \end{pmatrix}$$

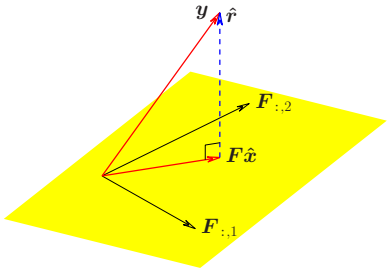
Vi vil forudsætte, at søjlerne i $\boldsymbol{F}$ er lineært uafhængige

Eksempel. $m = n$: $\boldsymbol{F}$ er en kvadratisk, regulær (iflg forudsætningen) matrix, og $\boldsymbol{r}(\boldsymbol{x}) = \boldsymbol{0}$ (og dermed $\varphi(\boldsymbol{x}) = 0$) for $\hat{\boldsymbol{x}} = \boldsymbol{F}^{-1}\boldsymbol{y}$, dvs løsningen til det lineære ligningssytem $\boldsymbol{F}\boldsymbol{x} = \boldsymbol{y}$

Hvis $m > n$ kan vi normalt ikke finde en $\boldsymbol{x}$, som opfylder alle ligninger i det overbestemte system $\boldsymbol{F}\boldsymbol{x} = \boldsymbol{y}$

Eksempel 4.3 – 4.4. $m = 3$, $n = 2$. Alle vektorer $\boldsymbol{z} = \boldsymbol{F}\boldsymbol{x}$ ligger i billedrummet for $\boldsymbol{F}$, som er planen udspændt af de to søjler i $\boldsymbol{F}$.

Pytagoras : $\varphi(\boldsymbol{x}) = \|\boldsymbol{r}(\boldsymbol{x})\|^2$ er mindst, når $\boldsymbol{r}$ er vinkelret på billedrummet



$\hat{\boldsymbol{r}} = \boldsymbol{y} - \boldsymbol{F}\hat{\boldsymbol{x}}$ ortogonal til alle søjler i $\boldsymbol{F}$: $\quad \boldsymbol{F}^T(\boldsymbol{y} - \boldsymbol{F}\hat{\boldsymbol{x}}) \;=\; \boldsymbol{0} \quad \Leftrightarrow$

Normalligninger : $\quad \left(\boldsymbol{F}^T\boldsymbol{F}\right)\hat{\boldsymbol{x}} \;=\; \boldsymbol{F}^T\boldsymbol{y}$

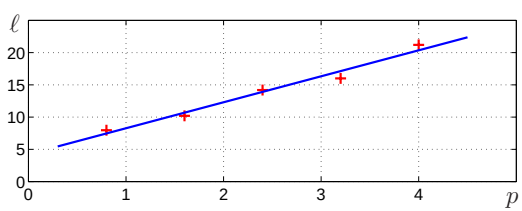
Tidligere vist: $n \times n$ matricen $\boldsymbol{A} = \boldsymbol{F}^T\boldsymbol{F}$ er SPD: symmetrisk og positivt definit.
Kan udnyttes til at vise, at løsningen til normalligningerne faktisk minimerer φ :

$$\begin{aligned}
\boldsymbol{r}(\hat{\boldsymbol{x}}+\boldsymbol{u}) \;&=\; \boldsymbol{y} - \boldsymbol{F}\,(\hat{\boldsymbol{x}}+\boldsymbol{u}) \;=\; \hat{\boldsymbol{r}} - \boldsymbol{F}\boldsymbol{u}\,. \qquad \boldsymbol{F}^T\hat{\boldsymbol{r}} = \boldsymbol{0} \\
\varphi(\hat{\boldsymbol{x}}+\boldsymbol{u}) \;&=\; (\hat{\boldsymbol{r}} - \boldsymbol{F}\boldsymbol{u})^T\,(\hat{\boldsymbol{r}} - \boldsymbol{F}\boldsymbol{u}) \\
&=\; \hat{\boldsymbol{r}}^T\hat{\boldsymbol{r}} - \hat{\boldsymbol{r}}^T\boldsymbol{F}\,\boldsymbol{u} - \boldsymbol{u}^T\boldsymbol{F}^T\hat{\boldsymbol{r}} + \boldsymbol{u}^T\boldsymbol{F}^T\boldsymbol{F}\,\boldsymbol{u} \\
&=\; \hat{\boldsymbol{r}}^T\hat{\boldsymbol{r}} + \boldsymbol{u}^T\boldsymbol{A}\,\boldsymbol{u} \;>\; \hat{\boldsymbol{r}}^T\hat{\boldsymbol{r}} \qquad \text{hvis } \boldsymbol{u} \neq \boldsymbol{0}
\end{aligned}$$

Eksempel 4.5. Fjeder

$$\boldsymbol{y} \;=\; \begin{pmatrix} 7.97 \\ 10.2 \\ 14.2 \\ 16.0 \\ 21.2 \end{pmatrix}, \qquad \boldsymbol{F} \;=\; \begin{pmatrix} 1 & 0.8 \\ 1 & 1.6 \\ 1 & 2.4 \\ 1 & 3.2 \\ 1 & 4.0 \end{pmatrix}, \qquad \begin{pmatrix} 5 & 12 \\ 12 & 35.2 \end{pmatrix}\begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} \;=\; \begin{pmatrix} 69.57 \\ 192.776 \end{pmatrix}, \qquad \hat{\boldsymbol{x}} \;=\; \begin{pmatrix} 4.2360 \\ 4.0325 \end{pmatrix}$$

Dvs $M(\hat{\boldsymbol{x}}, p) = 4.2360 + 4.0325p$ er den bedste model i mindste kvadraters forstand

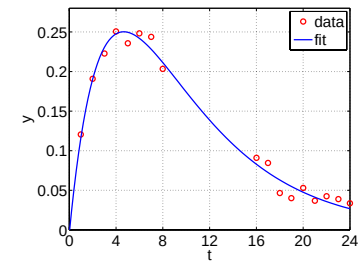


Eksempel 4.6. Koncentration af medicin i nyrerne.
Compartment model $\quad Y(t) \;=\; c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}$

Hvis vi kender λ_1 og λ_2 , er

$$M(\boldsymbol{x}, t) \;=\; x_1 e^{\lambda_1 t} + x_2 e^{\lambda_2 t}$$

en god fitte model. Fittet er beregnet med $\lambda_1 = -0.15,\ \lambda_2 = -0.30$, og mindste kvadraters løsningen er $\hat{x}_1 = 1.006$, $\hat{x}_2 = -1.011$



Hvis vi ikke kender λ_1 og λ_2 , kan vi bruge fitte modellen

$$M(\boldsymbol{x}, t) \;=\; x_1 e^{x_3 t} + x_2 e^{x_4 t}$$

Ulineær model. Bestemmelsen af $\hat{\boldsymbol{x}}$ er et af emnerne i kursus 02610 *Optimering og data fitting*

MATLAB: Data i søjle-vektorer `t` og `y` . `lam = [-0.15, -0.30]` (række-vektor)

```
F = exp(t*lam);   xh = F \ y
```

`xh` indeholder mindste kvadraters løsningen.
Hvordan beregnes den ? (Ikke via normalligningerne)

4. Lineære mindste kvadraters problemer pp 61 – 66

Givet m-vektoren $\boldsymbol{y}$ og $m \times n$-matricen $\boldsymbol{F}$ (med $n \leq m$). Søger $\hat{\boldsymbol{x}}$, den $\boldsymbol{x}$, som minimerer $\|\boldsymbol{y} - \boldsymbol{F}\boldsymbol{x}\|$. Samme vektor, som minimerer

$$\|\boldsymbol{y} - \boldsymbol{F}\boldsymbol{x}\|^2 \;=\; \|\boldsymbol{r}(\boldsymbol{x})\|^2 \;=\; \boldsymbol{r}(\boldsymbol{x})^T\boldsymbol{r}(\boldsymbol{x}) \;=\; r_1(\boldsymbol{x})^2 + r_2(\boldsymbol{x})^2 + \cdots + r_m(\boldsymbol{x})^2$$

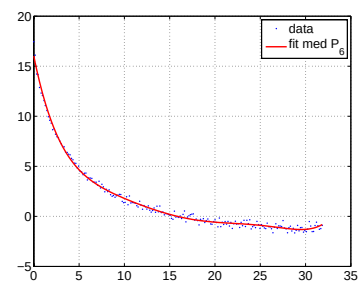
Eksempel. Data fitting.

Givet (måle)punkter $(t_1, y_1), (t_2, y_2), \ldots, (t_m, y_m)$
og lineær fitte model

$$\begin{aligned}
M(\boldsymbol{x}, t) \;&=\; x_1 f_1(t) + x_2 f_2(t) + \cdots + x_n f_n(t) \\
r_i(\boldsymbol{x}) \;&=\; y_i - M(\boldsymbol{x}, t_i) \;=\; y_i - \boldsymbol{F}_{i,:}\boldsymbol{x}\,,
\end{aligned}$$

idet $(\boldsymbol{F})_{ij} = f_j(t_i)$

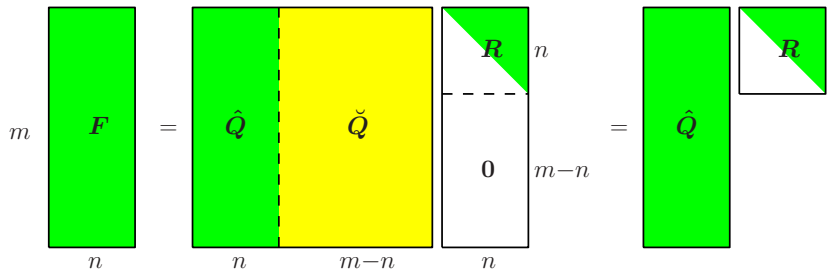
Figur: Lys-intensitet i optisk fiber efter afbrydelse af lyskilden. Fit med 6. grads polynomium,
$f_1(t) = 1,\ f_2(t) = t,\ \ldots,\ f_7(t) = t^6$



Mindste kvadraters løsningen kan findes via Normalligningerne : $\left(\boldsymbol{F}^T\boldsymbol{F}\right)\hat{\boldsymbol{x}} \;=\; \boldsymbol{F}^T\boldsymbol{y}$.
Mere nøjagtige beregninger via

4.2. QR faktorisering.

Kan bestemme en ortogonal matrix $\boldsymbol{Q} = \begin{pmatrix} \hat{\boldsymbol{Q}} & \breve{\boldsymbol{Q}} \end{pmatrix}$
så



Lineært uafhængige søjler i $\boldsymbol{F}$ $\Leftrightarrow$ alle $r_{ii} \neq 0$. Forudsættes

$\boldsymbol{F} = \hat{\boldsymbol{Q}}\boldsymbol{R}$ er økonomiudgaven af QR faktoriseringen.

Søjlerne i $\hat{\boldsymbol{Q}}$ er en ortonormal basis for billedrummet for $\boldsymbol{F}$

Vilkårlig ortogonal $\boldsymbol{Q}$: $\|\boldsymbol{Q}\boldsymbol{r}\|^2 \;=\; \boldsymbol{r}^T\boldsymbol{Q}^T\boldsymbol{Q}\,\boldsymbol{r} \;=\; \boldsymbol{r}^T\boldsymbol{r} \;=\; \|\boldsymbol{r}\|^2$

Dvs $\min\|\boldsymbol{r}(\boldsymbol{x})\|$ for samme $\boldsymbol{x}$ som $\min\|\boldsymbol{Q}\boldsymbol{r}(\boldsymbol{x})\|$

Brug $\boldsymbol{Q}^T$ fra QR faktoriseringen: $\boldsymbol{F} \;=\; \boldsymbol{Q}\begin{pmatrix}\boldsymbol{R}\\ \boldsymbol{0}\end{pmatrix} \quad\Rightarrow\quad \boldsymbol{Q}^T\boldsymbol{F} \;=\; \begin{pmatrix}\boldsymbol{R}\\ \boldsymbol{0}\end{pmatrix}$

$$\boldsymbol{Q}^T\boldsymbol{r}(\boldsymbol{x}) \;=\; \boldsymbol{Q}^T(\boldsymbol{y}-\boldsymbol{F}\boldsymbol{x}) \;=\; \boldsymbol{Q}^T\boldsymbol{y} - \begin{pmatrix}\boldsymbol{R}\\ \boldsymbol{0}\end{pmatrix}\boldsymbol{x} \;=\; \begin{pmatrix}\hat{\boldsymbol{Q}}^T\boldsymbol{y}-\boldsymbol{R}\boldsymbol{x}\\ \breve{\boldsymbol{Q}}^T\boldsymbol{y}\end{pmatrix}$$

$$\left\|\begin{pmatrix}\hat{\boldsymbol{Q}}^T\boldsymbol{y}-\boldsymbol{R}\boldsymbol{x}\\ \breve{\boldsymbol{Q}}^T\boldsymbol{y}\end{pmatrix}\right\|^2 \;=\; \|\hat{\boldsymbol{Q}}^T\boldsymbol{y}-\boldsymbol{R}\boldsymbol{x}\|^2 + \|\breve{\boldsymbol{Q}}^T\boldsymbol{y}\|^2$$

Minimum $\|\boldsymbol{Q}^T\boldsymbol{r}(\boldsymbol{x})\| = \|\boldsymbol{r}(\boldsymbol{x})\|$ for $\hat{\boldsymbol{x}}$ bestemt ved $\boldsymbol{R}\hat{\boldsymbol{x}} \;=\; \hat{\boldsymbol{Q}}^T\boldsymbol{y}$
Hvis forudsætningen er opfyldt, er $\boldsymbol{R}$ regulær, og $\hat{\boldsymbol{x}}$ éntydig

Eksempel 4.7. MATLAB

```
x = F \ y
```

beregner løsningen via QR faktorisering af $\boldsymbol{F}$

```
[Q, RR] = qr(F)
```

returnerer hele $\boldsymbol{Q}$ og $m \times n$ matricen $\begin{pmatrix}\boldsymbol{R}\\ \boldsymbol{0}\end{pmatrix}$. `R = RR(1:n,:)` giver $\boldsymbol{R}$

```
[Qh, R] = qr(F,0)
```

giver økonomiudgaven $\hat{\boldsymbol{Q}}$ og $\boldsymbol{R}$