Fitting Data by Least Squares

Algorithms and Examples

Walter Gander

Institute for Computational Science

Computer Science Departement

ETH Zürich

4 CiiT 2003

Bitola

11th - 14th Dec 2003

Chemical Reaction

- Model: decrease of component in chemical reaction

$$f(t) = a_0 + a_1 e^{-bt}$$

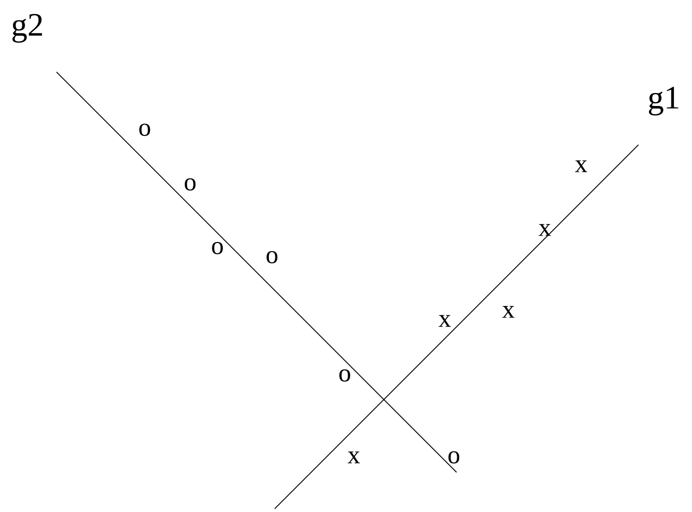
- Data: measure amount of component for different times

t	t_1	$\cdots$	t_m
f	f_1	$\cdots$	f_m

- Fit Parameters: estimate a_0, a_1 and b from measured data

Nonlinear Least squares problem

Coordinate Metrology



Fit orthogonal lines by minimizing the geometric distance

$$\begin{aligned} g_1: \quad c_1 + n_1 x + n_2 y &= 0 \\ g_2: \quad c_2 - n_2 x + n_1 y &= 0 \end{aligned} \qquad n_1^2 + n_2^2 = 1$$

Inserting measured points we get

Fitting orthogonal lines

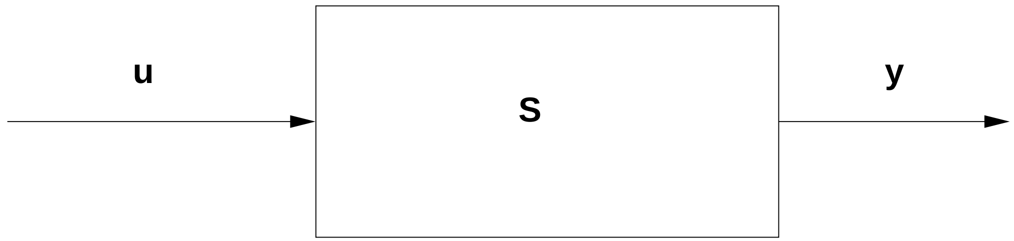
$$\begin{pmatrix} 1 & 0 & x_{P_1} & y_{P_1} \\ 1 & 0 & x_{P_2} & y_{P_2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & x_{P_p} & y_{P_p} \\ 0 & 1 & y_{Q_1} & -x_{Q_1} \\ 0 & 1 & y_{Q_2} & -x_{Q_2} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & y_{Q_q} & -x_{Q_q} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ n_1 \\ n_2 \end{pmatrix} \approx \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{s.t. } n_1^2 + n_2^2 = 1$$

Linear least squares with nonlinear constraint

parallel

Control Theory

System with input $\mathbf{u}$ and output $\mathbf{y}$



linear model

$$y_{t+n} + a_{n-1}y_{t+n-1} + \cdots + a_0 y_t \approx b_{n-1}u_{t+n-1} + b_{n-2}u_{t+n-2} + \cdots + b_0 u_t$$

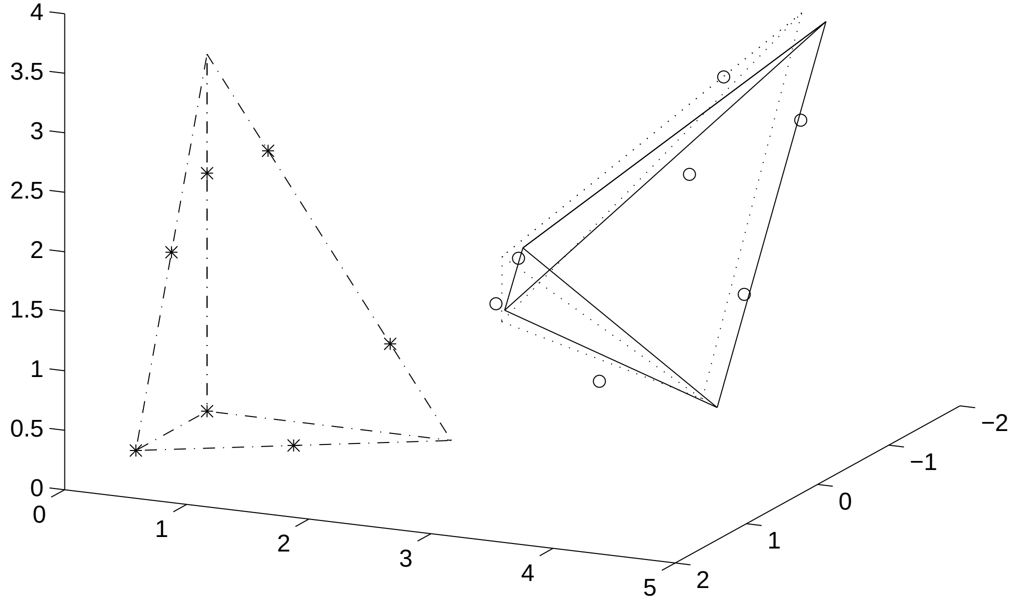
New equation for each time step

Control Theory (cont.)

$$\begin{bmatrix} y_{n-1} & y_{n-2} & \cdots & y_0 & -u_{n-1} & -u_{n-2} & \cdots & -u_0 \\ y_n & y_{n-1} & \cdots & y_1 & -u_n & -u_{n-1} & \cdots & -u_1 \\ y_{n+1} & y_n & \cdots & y_2 & -u_{n+1} & -u_n & \cdots & -u_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} a_{n-1} \\ a_{n-2} \\ \vdots \\ a_0 \\ b_{n-1} \\ b_{n-2} \\ \vdots \\ b_n \end{bmatrix} \approx \begin{bmatrix} -y_n \\ -y_{n+1} \\ -y_{n+2} \\ \vdots \end{bmatrix}$$

- linear least squares problem
- Add/subtract equations: up- downdate techniques
- structured: Toeplitz matrices

Robotics



Move body

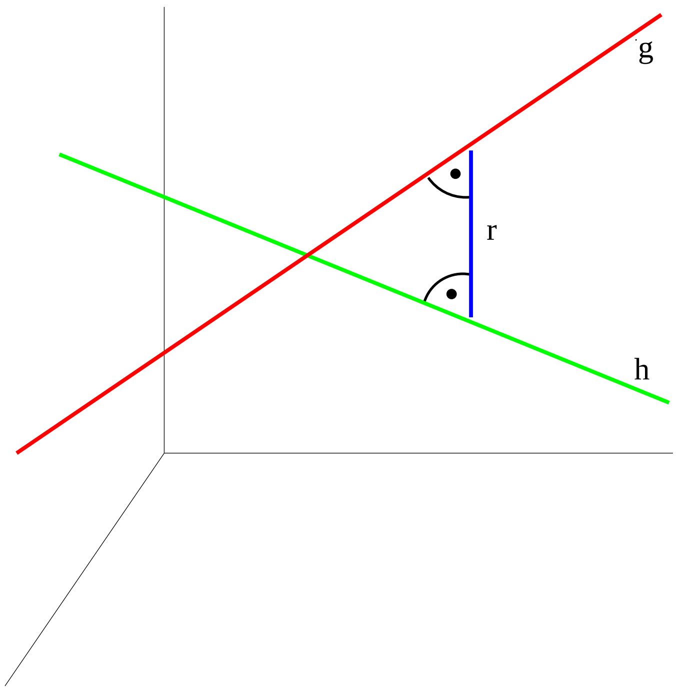
measured points $\boldsymbol{\xi}_i$

nominal points $\mathbf{x}_i$

Procrustes problem: Find Q orthogonal (3 rotations) and translation $\mathbf{t}$ such that

$$\boldsymbol{\xi}_i \approx Q\mathbf{x}_i + \mathbf{t} \quad i = 1, \ldots, m$$

Warped Lines



- given 2 lines in space

$$g: \quad \mathbf{X} = \mathbf{P} + \lambda \mathbf{r}$$
$$h: \quad \mathbf{Y} = \mathbf{Q} + \mu \mathbf{s}$$

- cutting point? $\quad \mathbf{P} + \lambda \mathbf{r} = \mathbf{Q} + \mu \mathbf{s}$

$$\begin{pmatrix} r_1 & -s_1 \\ r_2 & -s_2 \\ r_3 & -s_3 \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} Q_1 - P_1 \\ Q_2 - P_2 \\ Q_3 - P_3 \end{pmatrix}$$

- solve as least squares problem:
  1. residual $= 0 \Rightarrow$ solution is cutting point
  2. residual $\neq 0 \Rightarrow$ warped: solution are nearest points

Least squares principle

- Given $\mathbf{f} : \mathbb{R}^n \mapsto \mathbb{R}^m$
- We want $\mathbf{x} \in \mathbb{R}^n$ with $f_i(x_1, x_2, \ldots, x_n) \approx 0, \quad i = 1, \ldots, m$

$$\Longleftrightarrow \quad \|\mathbf{f}(\mathbf{x})\| = \min \quad \text{for some norm}$$

- For 2-norm $\quad \|\mathbf{f}(\mathbf{x})\|_2^2 = \sum_{i=1}^m f_i(\mathbf{x})^2 = \min$

Least Squares 1795

C.F. Gauss 1777–1855

Data Mining by Gauss

- Data: June 1801, Zach publishes orbital positions of Ceres discovered by G. Piazzi January 1, 1801 (only observations of 9 degrees before disappearing behind the sun).
- Mining techniques: Zach published several predictions of Ceres' position, including one by Gauss which differed greatly from the others.
- Success of Least Squares: When Ceres was rediscovered by Zach on December 7, 1801 it was almost exactly where Gauss had predicted by his least squares method.

Newton's method for n equations with n unknowns

- $\mathbf{f}(\mathbf{x}) : \mathbb{R}^n \mapsto \mathbb{R}^n$ nonlinear function
- Zero of $\mathbf{f}$: find $\mathbf{x}$ such that $\mathbf{f}(\mathbf{x}) = 0$
- linearize: $\mathbf{f}(\mathbf{x}) \approx \mathbf{f}(\mathbf{x}_k) + J(\mathbf{x}_k)\mathbf{h} = 0$ with $\mathbf{h} = \mathbf{x} - \mathbf{x}_k$

  Jacobian matrix: $$J = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$
- iterate: $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{h} = \mathbf{x}_k - J(\mathbf{x}_k)^{-1}\mathbf{f}\ (\mathbf{x}_k)$

Nonlinear least squares

- $\mathbf{f}(\mathbf{x}) \approx 0$ where $\mathbf{f}(\mathbf{x}) : \mathbb{R}^n \mapsto \mathbb{R}^m$ with $m > n$
- Gauss principle:

$$\Phi(\mathbf{x}) := f_1(\mathbf{x})^2 + f_2(\mathbf{x})^2 + \cdots f_m(\mathbf{x})^2 = \min$$

- necessary: $\operatorname{grad} \Phi = 0 \Rightarrow$ solve by Newton's method
- $\frac{\partial \Phi(\mathbf{x})}{\partial x_i} = \sum_{l=1}^m f_l(\mathbf{x}) \frac{\partial f_l}{\partial x_i}, \iff \operatorname{grad} \Phi(\mathbf{x}) = J(\mathbf{x})^T \mathbf{f}(\mathbf{x})$
- Jacobian of $\operatorname{grad} \Phi(\mathbf{x})$ = Hessian of $\Phi(\mathbf{x})$
- Newton correction: solve $\operatorname{hess} \Phi(\mathbf{x}_k)\, \mathbf{h} = -J(\mathbf{x}_k)^T \mathbf{f}(\mathbf{x}_k)$

$\Rightarrow \frac{\partial^2 \Phi(\mathbf{x})}{\partial x_i \partial x_j} = \sum_{l=1}^m \frac{\partial f_l}{\partial x_j} \frac{\partial f_l}{\partial x_i} + \sum_{l=1}^m f_l(\mathbf{x}) \frac{\partial^2 f_l}{\partial x_i \partial x_j}$

$$\left(J(\mathbf{x}_k)^T J(\mathbf{x}_k) + \sum_{l=1}^m f_l(\mathbf{x}_k) \operatorname{hess} f_l(\mathbf{x}_k) \right) \mathbf{h} = -J(\mathbf{x}_k)^T \mathbf{f}(\mathbf{x}_k)$$

Variants of nonlinear least squares

$$\left(J(\mathbf{x}_k)^T J(\mathbf{x}_k) + \sum_{l=1}^{m} f_l(\mathbf{x}_k)\,\text{hess}\, f_l(\mathbf{x}_k)\right)\mathbf{h} = -J(\mathbf{x}_k)^T\mathbf{f}(\mathbf{x}_k)$$

Simplification: Levenberg-Marquard or Tikhonov regularization

$$\left(J(\mathbf{x}_k)^T J(\mathbf{x}_k) + \tau^2 D^2\right)\mathbf{h} = -J(\mathbf{x}_k)^T\mathbf{f}(\mathbf{x}_k) \iff \begin{pmatrix} J(\mathbf{x}_k) \\ \tau D \end{pmatrix}\mathbf{h} \approx \begin{pmatrix} \mathbf{f}(\mathbf{x}_k) \\ 0 \end{pmatrix}$$

$$\iff \text{minimize } ||J(\mathbf{x}_k)\mathbf{h} - \mathbf{f}(\mathbf{x}_k||^2 + \tau^2||D\mathbf{h}||^2$$

- $D = 0$: Gauss-Newton Method $J(\mathbf{x}_k)\mathbf{h} \approx \mathbf{f}(\mathbf{x}_k)$
- $D \neq 0$: Trust region method

$$\iff \min ||J(\mathbf{x}_k)\mathbf{h} - \mathbf{f}(\mathbf{x}_k)||^2 \quad s.t. \quad ||D\mathbf{h}||^2 \leq \alpha^2$$

Example 1 $\quad f_l(\mathbf{x}) = x_1 + x_2 e^{-x_3 t_l} - y_l$

$$\operatorname{grad} f_l = \begin{pmatrix} 1 \\ e^{-x_3 t_l} \\ -x_2 t_l e^{-x_3 t_l} \end{pmatrix} \quad \operatorname{hess} f_l = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_l e^{-x_3 t_l} \\ 0 & -t_l e^{-x_3 t_l} & x_2 t_l^2 e^{-x_3 t_l} \end{pmatrix}$$

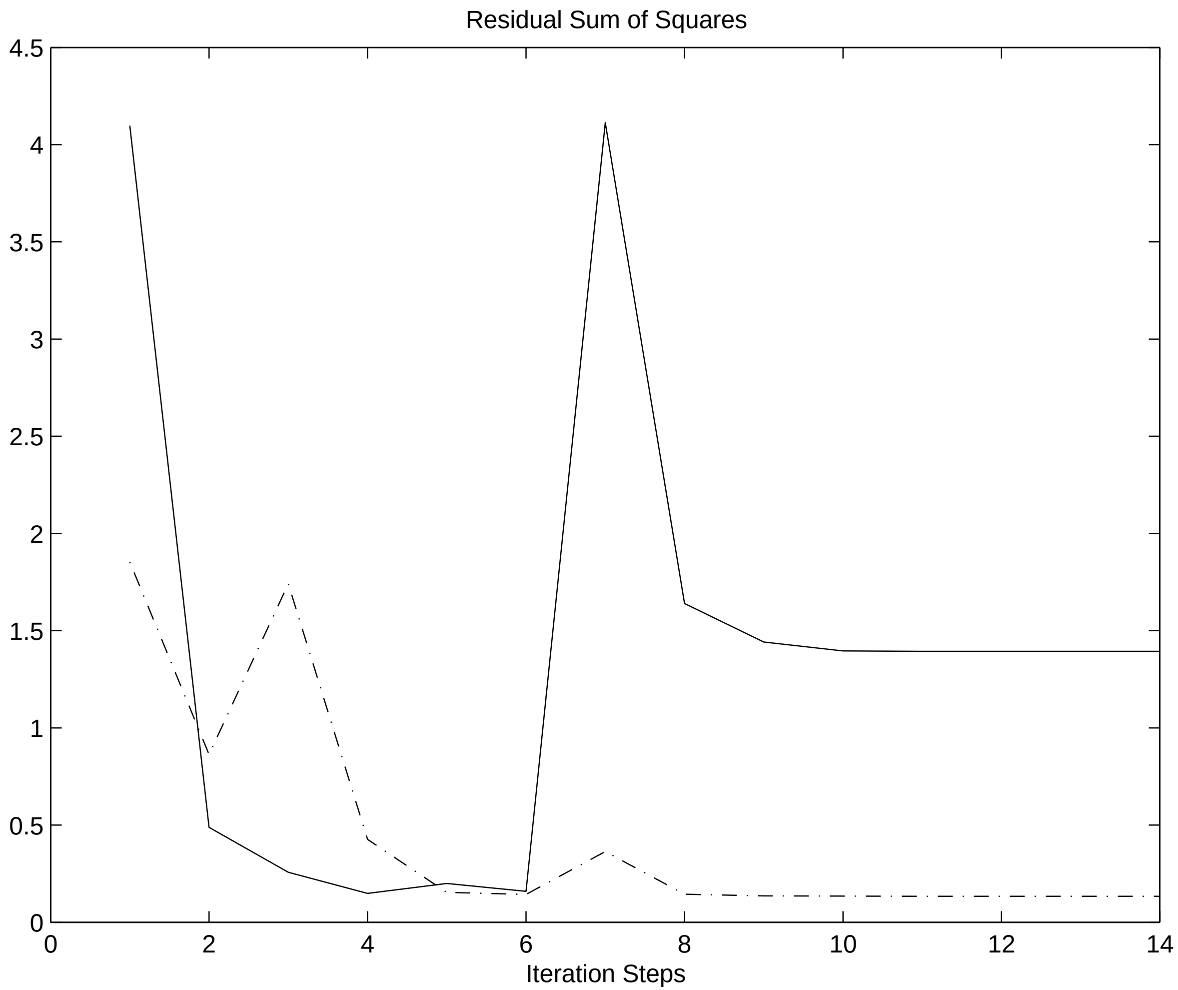
```
function  [x,fv] = newton(x,xi,eta)
h=x; fv=[]
while (norm(h)> 100*eps*norm(x)),
  ee = exp(-x(3)*xi);  tee =  xi.*ee;
  J = [ones(size(xi)) ee -x(2)*tee ];  % Jacobian
  f = x(1) + x(2)*ee - eta;            % residual
  s1 = f'*tee;  s2 = x(2)*f'*(xi.*tee);
  A = J'*J + [0 0 0; 0 0 -s1; 0 -s1 s2]; % Hessian of Phi
  h = -A\(J'*f);  x = x + h;
  [x' norm(h) norm(J'*f)],  fv=[fv norm(f)];
end
```

Example 1 (cont.)

```
% define the exact function:
  a0 = 1;   a1 = 2;   b = 0.15;
% and the exact function values
  xi = [1:0.3:7]';  etae = a0 + a1*exp(-b*xi);
% now compute ``measured values'' by adding a  small
% perturbation to the exact function values
  rand('seed',0);
  eta = etae + 0.1*(rand(size(etae))-0.5);
% first computation
  [x1,fv1]= newton([1.8 1.8 0.1]',xi,eta);
% second with different starting values
  [x2,fv2]= newton([1.5 1.5 0.1]',xi,eta);
```

Results $x1 = [2.1366,\ 0.0000,\ 0.0000]$ $x2 = [1.1481,\ 1.8623,\ 0.1702]$

```
plot([1:14], fv1(1:14),'-',[1:14], fv2(1:14),':')
```



Linear Least Squares

$$A \in \mathbb{R}^{m \times n}, \qquad \|\mathbf{b} - A\mathbf{x}\|_2 = \min \iff A\mathbf{x} \approx \mathbf{b}$$

Note: if $A \in \mathbb{R}^{n \times n}$ then

$$A\mathbf{x} = \mathbf{b} \iff BA\mathbf{x} = B\mathbf{b} \quad \text{if } B \text{ is nonsingular}$$

However,

$$A\mathbf{x} \approx \mathbf{b} \iff BA\mathbf{x} \approx B\mathbf{b} \quad \text{if } B \text{ is orthogonal :} \quad B^T B = I$$

Therefore: orthogonal matrices important for least squares

Solution with QR

QR-decomposition

$$A = Q\begin{pmatrix} R \\ 0 \end{pmatrix} \quad \text{with } Q \text{ orthogonal}$$

Equivalent system

$$Q^T A\mathbf{x} = \begin{pmatrix} R \\ 0 \end{pmatrix}\mathbf{x} \approx \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}, \quad \text{with } \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} := Q^T\mathbf{b}$$

Residual

$$||\mathbf{r}||^2 = ||\mathbf{y}_1 - R\mathbf{x}||^2 + ||\mathbf{y}_2||^2$$

is minimal for

$$R\hat{\mathbf{x}} = \mathbf{y}_1, \quad \hat{\mathbf{x}} = R^{-1}\mathbf{y}_1 \quad \text{and} \quad \min||\mathbf{r}|| = ||\mathbf{y}_2||$$

Normal equations $A^T A\mathbf{x} = A^T b$: condition number is squared

Singular value decomposition

$A \in \mathbb{R}^{m\times n}$ with $m \geq n$ there exist

- $U \in \mathbb{R}^{m\times n}$ and $V \in \mathbb{R}^{n\times n}$ orthogonal, $U^T U = V^T V = I$
- $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^{m\times n}$ diagonal such that

$$A = U\Sigma V^T \qquad \text{singular value decomposition}$$

- Singular values: $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$
- Rank: if $\sigma_r > 0$ and $\sigma_{r+1} = \cdots \sigma_n = 0$ then A has rank r
- Equivalent system with diagonal matrix

$$A\mathbf{x} \approx \mathbf{b} \iff \Sigma\mathbf{y} \approx \mathbf{c} \quad \text{with} \quad \mathbf{y} = V^T\mathbf{x}, \quad \mathbf{c} = U^T\mathbf{b}$$

General solution of $A\mathbf{x} \approx \mathbf{b}$ with SVD

1. Compute the SVD: `[U S V] = svd(A)`
2. Make a rank decision, choose r with $\sigma_r > 0$ and $\sigma_{r+1} = \cdots = \sigma_n = 0$ (necessary because of rounding errors)
3. Partition
   `V1=V[:,1:r], V2= V[:,r+1:n], Sr=S[1:r,1:r], U1=U[:,1:r]`
4. Minimal norm solution `xm=V1*(Sr\U1'*b)` $= A^+\mathbf{b}$
5. General solution `x = xm + V2*c` with arbitrary $\mathbf{c} \in \mathbb{R}^{n-r}$ same as

$$\mathbf{x} = A^+\mathbf{b} + (I - A^+A)\mathbf{w}, \quad \mathbf{w} \text{ arbitrary}$$

6. Projectors

   1. $P_{\mathcal{R}(A)} = AA^+ = U_1U_1^T$
   2. $P_{\mathcal{R}(A^T)} = A^+A = V_1V_1^T$
   3. $P_{\mathcal{N}(A^T)} = I - AA^+ = U_2U_2^T$
   4. $P_{\mathcal{N}(A)} = I - A^+A = V_2V_2^T$

Some quadratically constraint problems

If $A = U\Sigma V^T$ and $V = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ then

$$\|A\mathbf{x}\|_2 = \min, \quad \text{subject to } \|\mathbf{x}\|_2 = 1$$

has the solution

$$\mathbf{x} = \mathbf{v}_n \quad \text{with} \quad \min_{\|\mathbf{x}\|_2=1} \|A\mathbf{x}\|_2 = \|A\mathbf{v}_n\|_2 = \sigma_n$$

Applications:

- Fitting Lines: $c + n_1 x + n_2 y = 0, \quad n_1^2 + n_2^2 = 1$
- Fitting Ellipses: $\mathbf{x}^T A\mathbf{x} + \mathbf{b}^T\mathbf{x} + c = 0$
  - coefficients: $\mathbf{u} = [a_{11}, a_{12}, a_{22}, b_1, b_2, c]$ since $A = A^T$
  - normalizing condition: $||\mathbf{u}||_2 = 1$

Fitting lines

- Equation $c + n_1 x + n_2 y = 0, \quad n_1^2 + n_2^2 = 1$
- Normal vector $\mathbf{n} = (n_1, n_2)^T$
- Distance of $P = (x_i, y_i)$ from line

$$d_i = |r|, \quad r = c + n_1 x_i + n_2 y_i$$

- Minimize geometric distances $\sum_{i=1}^{m} d_i^2 = \min$

$$\begin{pmatrix} 1 & x_{P_1} & y_{P_1} \\ 1 & x_{P_2} & y_{P_2} \\ \vdots & \vdots & \vdots \\ 1 & x_{P_m} & y_{P_m} \end{pmatrix} \begin{pmatrix} c \\ n_1 \\ n_2 \end{pmatrix} \approx \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{subject to } n_1^2 + n_2^2 = 1$$

Fitting lines (cont.)

$$A\mathbf{x} \approx 0, \quad \text{subject to } n_1^2 + n_2^2 = 1 \quad \mathbf{x} = (c, n_1, n_2)^T$$

QR decomposition $A\mathbf{x} = Q\binom{R}{0}\mathbf{x} \approx 0 \iff R\mathbf{x} \approx 0$

$$\begin{pmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{pmatrix} \begin{pmatrix} c \\ n_1 \\ n_2 \end{pmatrix} \approx \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{s.t. } n_1^2 + n_2^2 = 1$$

$$\iff \begin{pmatrix} r_{22} & r_{23} \\ 0 & r_{33} \end{pmatrix} \begin{pmatrix} n_1 \\ n_2 \end{pmatrix} \approx \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{s.t. } n_1^2 + n_2^2 = 1$$

Solve by SVD

$$\mathbf{n} = \mathbf{v}_2$$

Two Parallel Lines

$$c_1 + n_1 x + n_2 y = 0, \quad c_2 + n_1 x + n_2 y = 0, \quad n_1^2 + n_2^2 = 1$$

$$\begin{pmatrix} 1 & 0 & x_{P_1} & y_{P_1} \\ 1 & 0 & x_{P_2} & y_{P_2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & x_{P_p} & y_{P_p} \\ 0 & 1 & x_{Q_1} & y_{Q_1} \\ 0 & 1 & x_{Q_2} & y_{Q_2} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & x_{Q_q} & y_{Q_q} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ n_1 \\ n_2 \end{pmatrix} \approx 0 \quad \text{subject to} \quad n_1^2 + n_2^2 = 1.$$

orth. lines

Fitting a Rectangle

$$
\begin{array}{rrcl}
a: & c_1 + n_1 x + n_2 y & = & 0 \\
b: & c_2 - n_2 x + n_1 y & = & 0 \\
c: & c_3 + n_1 x + n_2 y & = & 0 \\
d: & c_4 - n_2 x + n_1 y & = & 0 \\
 & n_1^2 + n_2^2 & = & 1
\end{array}
\qquad
\begin{pmatrix}
1 & 0 & 0 & 0 & x_{P_1} & y_{P_1} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & 0 & 0 & 0 & x_{P_p} & y_{P_p} \\
0 & 1 & 0 & 0 & y_{Q_1} & -x_{Q_1} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 1 & 0 & 0 & y_{Q_q} & -x_{Q_q} \\
0 & 0 & 1 & 0 & x_{R_1} & y_{R_1} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 1 & 0 & x_{R_r} & y_{R_r} \\
0 & 0 & 0 & 1 & y_{S_1} & -x_{S_1} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 1 & y_{S_s} & -x_{S_s}
\end{pmatrix}
\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ n_1 \\ n_2 \end{pmatrix}
\approx 0 \quad \text{s.t. } n_1^2 + n_2^2 = 1
$$

Fitting Ellipses

- Minimizing the algebraic distance

$$\mathbf{x}^T A\mathbf{x} + \mathbf{b}^T\mathbf{x} + c = 0$$

- unknown coefficients $\mathbf{u} = [a_{11}, a_{12}, a_{22}, b_1, b_2, c]$
- normalize the coefficients by $\|\mathbf{u}\|_2 = 1$
- Inserting measured points $\mathbf{x}_i^T A\mathbf{x}_i + \mathbf{b}^T\mathbf{x}_i + c \approx 0$

$$||B\mathbf{u}|| = \min \quad \text{subject to } ||\mathbf{u}||_2 = 1$$

Fitting Ellipses (cont.)

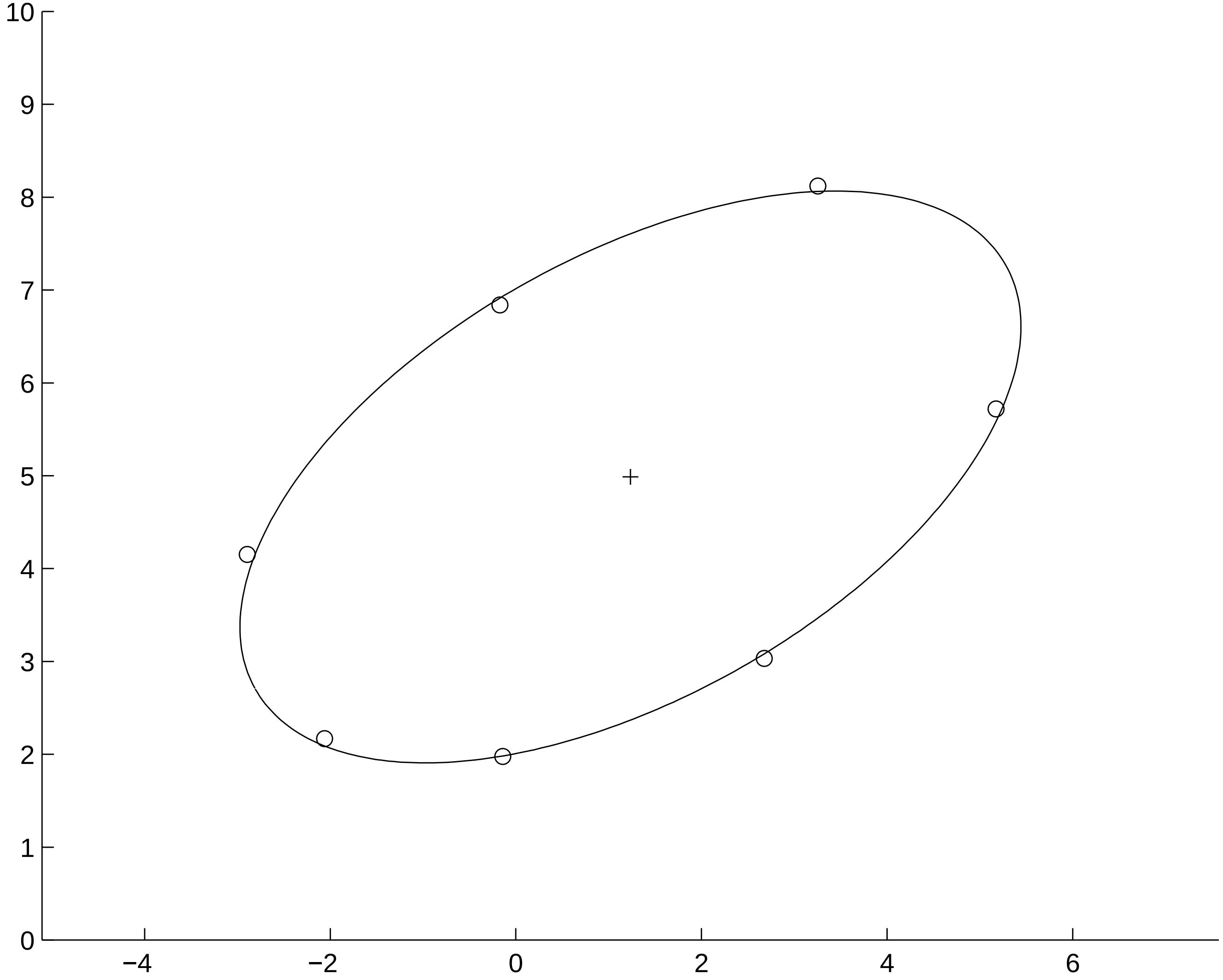
```
X =[-2.8939    4.1521
       ...        ...
    -0.1724    6.8398 ]
B =  [X(:,1).^2  X(:,1).*X(:,2) X(:,2).^2 ...
           X(:,1) X(:,2) ones(size(X(:,1)))]
```

Solution by SVD:

```
[U S V]= svd(B);  u = V(:,6);
A = [u(1) u(2)/2; u(2)/2 u(3)]; b  = [u(4); u(5)]; c = u(6);
```

$$A = \begin{pmatrix} -0.0316 & 0.0227 \\ 0.0227 & -0.0589 \end{pmatrix} \qquad b = \begin{pmatrix} -0.1484 \\ 0.5316 \end{pmatrix} \qquad c = -0.8300$$

Fitting Ellipses (cont.)



Least Squares Fit of Point Clouds

pyramid

- given m points $\mathbf{x}_i$ of workpiece in nominal position
- measure $\boldsymbol{\xi}_i$ in another frame of reference
- Problem: determine frame transformation

$$\boldsymbol{\xi}_i \approx Q\mathbf{x}_i + \mathbf{t} \iff \sum_{i=1}^{m} \|Q\mathbf{x}_i + \mathbf{t} - \boldsymbol{\xi}_i\|^2 = \min$$

Solution:

1. Compute *centers of gravity*: $\bar{\boldsymbol{\xi}} = \frac{1}{m}\sum_{i=1}^{m} \boldsymbol{\xi}_i, \quad \bar{\mathbf{x}} = \frac{1}{m}\sum_{i=1}^{m} \mathbf{x}_i$
2. Compute *relative coordinates*: $\mathbf{a}_i := \mathbf{x}_i - \bar{\mathbf{x}}$, $\mathbf{b}_i := \boldsymbol{\xi}_i - \bar{\boldsymbol{\xi}}$

   $A = [\mathbf{a}_i]$, $B = [\mathbf{b}_i]$
3. Procrustes problem $\|B^T - A^T Q^T\|_F^2 = \min$

   solved with SVD: $AB^T = U\Sigma V^T \Rightarrow Q = VU^T$, $\mathbf{t} = \bar{\boldsymbol{\xi}} - Q\bar{\mathbf{x}}$

References

1. Åke Björck, Numerical Methods for Least Squares Problems, SIAM, 1996.

2. W. Gander and J. Hřebíček, Solving Problems in Scientific Computing using Maple and Matlab, Springer Verlag, third edition 1997.

3. W. Gander, G. H. Golub and R. Strebel, Least-Squares Fitting of Circles and Ellipses, BIT 34, 1994, pp. 558-578.