Numerical analysis

Chapter 4: Least squares

Least squares and normal equations

Used when no solution can be found for a system of equations

Inconsistent systems

Systems without solutions

Normal equation

$A^T A \bar{x} = A^T b$ where $\bar{x}$ is the approx. solution that minimizes the distance

Ways to express the residual size

Euclidean length / 2-norm

$$\|r\|_2 = \sqrt{r_1^2 + \dots + r_m^2}$$

Squared error

$$SE = r_1^2 + \dots + r_m^2$$

Root mean squared error

$$RMSE = \sqrt{\frac{SE}{m}}$$

Fitting models to data

1. Choose model
2. Force model to fit data
3. Solve normal equations

Conditioning of least squares

Honestly not sure what is being explained

A survey of models

Periodic models

Natural systems often call for cosine or sine models

Data linearization

An exponential model cannot directly be expressed linearly, so we linearize it.

For example, by applying the natural logarithm to it

QR factorization

Gram-Schmidt

Orthagonalizes a given matrix
Each column of Q is calculated
by subtracting the prior column q_{n-1}
multiplied by q_{n-1}^T and A_n
from A_n

Diagonal elements of R are computed
by calculating the corresponding columns
length

$r_{nn} = \| A_n \|$

and remaining ones:

$r_{nm} = q_n^T A_m$

(see book for algo.)

def 4.1 A square matrix is orthogonal if $Q^T = Q^{-1}$

Least squares

1. QR - Factorize A
2. Solve $\hat{R}\bar{x} = \hat{d}$

(where $\hat{R}$ is upper triangular R and $\hat{d}$ is upper n entries of $Q^T b$)

Modified

(An enhanced variant for machines)

See algo. in book. A_m is computed in the outer for-loop

Householder reflectors

An orthogonal matrix that leaves vector lengths unchanged when multiplied with it - ideal for moving vectors.

Requires fewer iterations to calculate

the QR factorization

$$H = I - \frac{2vv^T}{v^Tv}$$

(Rest of chapter is not part of my course)

Exercises

4.1 Least squares and normal equations

1a) $$A^TA = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 2 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 4 \\ 4 & 6 \end{bmatrix}$$

$$A^Tb = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 8 \end{bmatrix}$$

In tableau form:

$$\left[\begin{array}{cc|c} 5 & 4 & 5 \\ 4 & 6 & 8 \end{array}\right] \Rightarrow$$ sub $\frac{4}{5} \times$ row 1 from row 2

$$\Rightarrow \left[\begin{array}{cc|c} 5 & 4 & 5 \\ 0 & \frac{14}{5} & 4 \end{array}\right]$$

$$\Rightarrow \bar{x}_2 = \frac{4 \cdot 5}{14} = \frac{20}{14} = \frac{10}{7}$$

$$\Rightarrow \bar{x}_1 = \frac{5 - 4 \cdot \frac{10}{7}}{5} = \frac{-\frac{5}{7}}{5} = -\frac{1}{7}$$

$$r = b - A\bar{x} = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{7} \\ \frac{10}{7} \end{bmatrix}$$

$$= \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} \frac{19}{7} \\ \frac{10}{7} \\ \frac{8}{7} \end{bmatrix} = \begin{bmatrix} \frac{2}{7} \\ -\frac{3}{7} \\ -\frac{1}{7} \end{bmatrix}$$

$$\|r\|_2 = \sqrt{\left(\frac{2}{7}\right)^2 + \left(-\frac{3}{7}\right)^2 + \left(-\frac{1}{7}\right)^2} = \sqrt{\frac{4+9+1}{49}}$$

$$= \frac{\sqrt{14}}{7}$$

b) $$A^T A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 14 & 6 \\ 6 & 3 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 14 & 6 & 5 \\ 6 & 3 & 3 \end{array}\right] \Rightarrow$$ sub $\frac{6}{14} \times$ row 1 from row 2

$$\Rightarrow \left[\begin{array}{cc|c} 14 & 6 & 5 \\ 0 & \frac{6}{14} & \frac{12}{14} \end{array}\right] \Rightarrow \bar{x}_2 = \frac{12}{6} = 2$$

$$\bar{x}_1 = \frac{5 - 12}{14} = -\frac{1}{2}$$

$$r = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} - \begin{bmatrix} \frac{3}{2} \\ 1 \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ 1 \\ -\frac{1}{2} \end{bmatrix}$$

$$\|r\|_2 = \sqrt{\frac{1}{4} + 1 + \frac{1}{4}} = \frac{\sqrt{6}}{2}$$

c)

$$A^T A = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 2 & 1 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 10 & 9 \\ 9 & 10 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 16 \\ 16 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 10 & 9 & 16 \\ 9 & 10 & 16 \end{array}\right] \Rightarrow$$ sub $\frac{9}{10}$ x row 1 from row 2

$$\Rightarrow \left[\begin{array}{cc|c} 10 & 9 & 16 \\ 0 & \frac{19}{100} & \frac{16}{100} \end{array}\right]$$

$$\Rightarrow \bar{x}_2 = \frac{16}{19}$$

$$\Rightarrow \bar{x}_1 = \frac{16 - 9 \cdot \frac{16}{19}}{10} = \frac{\frac{304 - 144}{19}}{10} = \frac{160}{190} = \frac{16}{19}$$

$$r = \begin{bmatrix} 3 \\ 3 \\ 3 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 2 & 1 \\ 2 & 2 \end{bmatrix} \frac{16}{19} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \\ 2 \end{bmatrix} - \frac{16}{19} \begin{bmatrix} 3 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

$$= \frac{1}{19}\begin{bmatrix} 57-48 \\ 57-32 \\ 57-48 \\ 38-64 \end{bmatrix} = \frac{1}{19}\begin{bmatrix} 9 \\ 25 \\ 9 \\ -26 \end{bmatrix}$$

$$\|r\|_2 = \frac{\sqrt{9^2+25^2+9^2+26^2}}{19} = \frac{\sqrt{1463}}{19} \approx 2.013$$

Skipping rest, this is all stuff we did in our linear algebra course :)

4.2 A survey of models

1a)

$F_3(t)$	y
$c_1 + c_2$	1
$c_1 \quad + c_3$	3
$c_1 - c_2$	2
$c_1 \quad - c_3$	0

corresponds to $Ac = y$

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \quad y = \begin{bmatrix} 1 \\ 3 \\ 2 \\ 0 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$A^T y = \begin{bmatrix} 6 \\ -1 \\ 3 \end{bmatrix} \quad \Rightarrow \quad \bar{c} = \begin{bmatrix} \frac{3}{2} \\ -\frac{1}{2} \\ \frac{3}{2} \end{bmatrix}$$

$$r = y - A\bar{c} = \begin{bmatrix} 1 \\ 3 \\ 2 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 3 \\ 2 \\ 0 \end{bmatrix} = 0$$

$$\|r\|_2 = 0 \Rightarrow RMSE = 0$$

b, c) shipping, tedious

2) -//-

3a) $\ln y = \ln(c_1 e^{c_2 t}) = \ln c_1 + c_2 t$

$$\begin{aligned} \ln c_1 - 2c_2 &= \ln 7 \\ \ln c_1 &= \ln 2 \\ \ln c_1 + c_2 &= \ln 2 \\ \ln c_1 + 2c_2 &= \ln 5 \end{aligned}$$

$$k = \ln c_1 \Rightarrow x = (k, c_2)$$

$$A = \begin{bmatrix} 1 & -2 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \quad y = \begin{bmatrix} \ln 1 \\ \ln 2 \\ \ln 2 \\ \ln 5 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 4 & 1 \\ 1 & 9 \end{bmatrix}$$

$$A^T y = \begin{bmatrix} \ln 1 + \ln 2 + \ln 2 + \ln 5 \\ -2\ln 1 + \ln 2 + 2\ln 5 \end{bmatrix}$$

$$= \begin{bmatrix} 2\ln 2 + \ln 5 \\ \ln 2 + 2\ln 5 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 4 & 1 & 2\ln 2 + \ln 5 \\ 1 & 9 & \ln 2 + 2\ln 5 \end{array}\right]$$

$\Rightarrow$ sub $\frac{1}{4}$ x row 1
from row 2

$$\Rightarrow \left[\begin{array}{cc|c} 4 & 1 & 2\ln 2 + \ln 5 \\ 0 & \frac{35}{4} & \frac{1}{2}\ln 2 + \frac{7}{4}\ln 5 \end{array}\right]$$

$$\Rightarrow c_2 = \frac{2\ln 2 + 7\ln 5}{35} \approx 0{,}361496$$

$$h = \frac{2\ln 2 + \ln 5 - c_2}{4} \approx 0,65856$$

$$\Rightarrow c_1 = e^h \approx 1,932$$

$$\Rightarrow f(t) = 1,932e^{0,36515t}$$

$$r = y - f(t)$$

$$= \begin{bmatrix} 1 \\ 2 \\ 2 \\ 5 \end{bmatrix} - \begin{bmatrix} 0.938 \\ 1,932 \\ 2,773 \\ 3,981 \end{bmatrix} = \begin{bmatrix} 0.062 \\ 0.068 \\ -0.773 \\ 1.019 \end{bmatrix}$$

$\|r\|_2 = \sqrt{r_1^2 + \ldots + r_m^2} \approx 1.2823$ (close enough)

b)

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \quad y = \begin{bmatrix} \ln 1 \\ \ln 1 \\ \ln 2 \\ \ln 4 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 4 & 4 \\ 4 & 6 \end{bmatrix}$$

$$A^T y = \begin{bmatrix} \ln 2 + \ln 4 \\ \ln 2 + 2\ln 4 \end{bmatrix} = \begin{bmatrix} 3\ln 2 \\ 5\ln 2 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 4 & 4 & 3\ln 2 \\ 4 & 6 & 5\ln 2 \end{array}\right] \Rightarrow$$ sub $1 \times$ row 1 from row 2

$$\Rightarrow \left[\begin{array}{cc|c} 4 & 4 & 3\ln 2 \\ 0 & 2 & 2\ln 2 \end{array}\right] \Rightarrow c_2 = \ln 2$$

$$\Rightarrow h = \frac{3\ln 2 - 4\ln 2}{4} = -\frac{1}{4}\ln 2$$

$$c_1 = 2^{-\frac{1}{4}} \Rightarrow f(t) = 2^{-\frac{1}{4}} \cdot 2^t = 2^{t-\frac{1}{4}}$$

$$r = y - f(t) = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 4 \end{bmatrix} - \begin{bmatrix} 0.841 \\ 1,682 \\ 1,682 \\ 3,364 \end{bmatrix} = \begin{bmatrix} 0.159 \\ -0,682 \\ 0.318 \\ 0.636 \end{bmatrix}$$

$\|r\|_2 \approx 0,998$ (close enough)

4) shipping

5) $y = c_1 t^{c_2} \Rightarrow \ln y = \ln c_1 + c_2 \ln t$

$h = \ln c_1$

$$A = \begin{bmatrix} 1 & 0 \\ 1 & \ln 2 \\ 1 & \ln 3 \\ 1 & \ln 4 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 4 & 3.178 \\ 3.178 & 3.609 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} 2.485 \\ 0.480 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 4 & 3.178 & 2.485 \\ 3.178 & 3.609 & 0.480 \end{array}\right] \Rightarrow$$ sub $\frac{3.178}{4} \times$ row 1 from row 2

$$\Rightarrow \left[\begin{array}{cc|c} 4 & 3.178 & 2.485 \\ 0 & 1.084 & -1.494 \end{array}\right]$$

$$\Rightarrow c_2 = \frac{-1.494}{1.084} \approx -1.379$$

$$k = \frac{2.485 + 3.178 \cdot 1.379}{4} \approx 1.716$$

$$c_1 = e^{1.716} \approx 5.565$$

$$f(t) = 5.565\, t^{-1.379}$$

close enough :D

b) Shipping to save on time?

6) -11-

4.3 QR factorization

1a)

$$y_1 = A_1 \qquad r_{11} = \sqrt{4^2+3^2} = 5$$

$$q_1 = \begin{bmatrix} \frac{4}{5} \\ \frac{3}{5} \end{bmatrix}$$

$$y_2 = A_2 - q_1 q_1^T A_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \begin{bmatrix} \frac{4}{5} \\ \frac{3}{5} \end{bmatrix} \begin{bmatrix} \frac{4}{5} & \frac{3}{5} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \frac{3}{5} \begin{bmatrix} \frac{4}{5} \\ \frac{3}{5} \end{bmatrix} = \begin{bmatrix} -\frac{12}{25} \\ \frac{16}{25} \end{bmatrix}$$

$$r_{12} = \frac{3}{5}$$

$$q_2 = \frac{5}{4} \begin{bmatrix} -\frac{12}{25} \\ \frac{16}{25} \end{bmatrix} = \begin{bmatrix} -\frac{3}{5} \\ \frac{4}{5} \end{bmatrix}$$

$$r_{21} = q_2^T A_1 = \frac{-12}{5} + \frac{12}{5} = 0$$

$$r_{22} = q_2^T A_2 = \frac{4}{5}$$

$$Q = \frac{1}{5}\begin{bmatrix} 4 & -3 \\ 3 & 4 \end{bmatrix} \quad R = \frac{1}{5}\begin{bmatrix} 25 & 3 \\ 0 & 4 \end{bmatrix}$$

b) $\eta_1 = A_1 \quad r_{11} = \sqrt{2}$

$$q_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\eta_2 = A_2 - q_1 q_1^T A_2$$

$$= \begin{bmatrix} 2 \\ 1 \end{bmatrix} - \frac{3}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$r_{12} = q_1^T A_2 = \frac{3}{\sqrt{2}}$$

$$q_2 = \sqrt{2} \cdot \frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$r_{21} = q_2^T A_1 = 0$

$r_{22} = q_2^T A_2 = \frac{2-1}{\sqrt{2}} = \frac{1}{\sqrt{2}}$

$Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad R = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}$

c) $y_1 = A_1 \qquad r_{11} = 3$

$q_1 = \frac{1}{3} \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$

$r_{12} = q_1^T A_1 = \frac{1}{3} \cdot 9 = 3$

$y_2 = A_2 - r_{12} \cdot q_1 = \begin{bmatrix} -1 \\ -2 \\ -1 \end{bmatrix}$

$r_{22} = \sqrt{6}$

$q_2 = -\frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$

We add a third vector A_3 to achieve full QR factorization

This gives

$$\eta_3 = A_3 - q_2 q_2^T A_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \frac{1}{6}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$$= \frac{1}{6}\begin{bmatrix} 7 \\ 2 \\ 1 \end{bmatrix}$$

$$q_3 = \frac{6}{\sqrt{54}}\begin{bmatrix} 7 \\ 2 \\ 1 \end{bmatrix} = \frac{6}{3\sqrt{6}}\begin{bmatrix} 7 \\ 2 \\ 1 \end{bmatrix} = \frac{\sqrt{6}}{3}\begin{bmatrix} 7 \\ 2 \\ 1 \end{bmatrix}$$

$$Q = \frac{1}{3}\begin{bmatrix} 2 & -\frac{3}{\sqrt{6}} & 7\sqrt{6} \\ 1 & -\sqrt{6} & 2\sqrt{6} \\ 2 & -\frac{3}{\sqrt{6}} & \sqrt{6} \end{bmatrix} \quad R = \begin{bmatrix} 3 & 3 \\ 0 & \sqrt{6} \\ 0 & 0 \end{bmatrix}$$

Let's check $A = QR$

$$QR = \frac{1}{3}\begin{bmatrix} 6 & 3 \\ 3 & -3 \\ 6 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 2 & 1 \end{bmatrix} = A \quad \checkmark \; ! \; :)$$

d) Skipping

2) -//-

3) Don't see the point, skipping

4) -//-

5) $H = I - \frac{2vv^T}{v^Tv}$ where $v = w - x$

(for every column)

$x_1 = [4 \quad 3]$ $w_1 = [\|x_1\|_2 \quad 0] = [5 \quad 0]$

$v_1 = [1 \quad -3]$

can't be bothered, actually

not part of my course

Skipping rest $\ddot{\nabla}$