Solution for Review Problems Set 1

Problem 1. Normal Equations:

In this problem we are going to fit a parabola to the four data points given below:

Data: (t,b): (1,1); (2,4); (3,9); (4,15)

Model: b = x(1) + x(2) * t + x(3) * t^2

a. Write down the least squares problem in the matrix form given in class, explicitly presenting the matrix A and the vector b.

b. Convert this problem to the corresponding normal equation problem (Hint: this new problem should be 3x3 - square).

c. (only if you have time left) solve this problem by gaussian elimination.

Solution

a.

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \begin{bmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ 1 & t_3 & t_3^2 \\ 1 & t_4 & t_4^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ 1 & t_3 & t_3^2 \\ 1 & t_4 & t_4^2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix}$$

$$b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 9 \\ 15 \end{bmatrix}$$

b. Normal Equation: $A^T A x = A^T b$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 16 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 16 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 9 \\ 15 \end{bmatrix}$$

c.

$$\begin{bmatrix} 4 & 10 & 30 \\ 10 & 30 & 100 \\ 30 & 100 & 354 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 29 \\ 96 \\ 338 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 10 & 30 \\ 0 & 5 & 25 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 29 \\ 47/2 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 19/20 \\ 3/4 \end{bmatrix}$$

$$b = -3/4 + 19/20t + 3/4t^2$$

Problem 2. Orthogonal Transformations:

a. A matrix P is orthogonal if P'P is the identity matrix. Prove that orthogonal transformations preserve the 2-norm.

b. How can this property of orthogonal transformations be useful in solving the linear least squares minimization problem min $\| Ax - b \|_2$? A long answer is not required.

Solution

a.

$$\| Px \|_2 = \left[(Px)^T (Px) \right]^{\frac{1}{2}}$$

$$= \left[x^T P^T P x \right]^{\frac{1}{2}}$$

$$= \left[x^T x \right]^{\frac{1}{2}} = \| x \|_2$$

$$\implies \| Px \|_2 = \| x \|_2$$

b. This property allows one to formulate QR decomposition.

Problem 3. Finite Differences:

The finite difference formula for a second derivative is given by:

$f"(x) \sim [f(x-h) + f(x+h) - 2f(x)]/h^2$

a. What is the error in this formula (include both algorithm error and machine error)?

b. Given a machine precision of eps, determine the optimum value of h and the minimum possible error.

Solution

a.

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2!}f''(x) + \frac{h^3}{3!}f'''(x) + \frac{h^4}{4!}f''''(x) + \cdots$$

$$f(x-h) = f(x) - hf'(x) + \frac{h^2}{2!}f''(x) - \frac{h^3}{3!}f'''(x) + \frac{h^4}{4!}f''''(x) + \cdots$$

$$f"(x) \sim \frac{[f(x-h) + f(x+h) - 2f(x)]}{h^2}$$

$$f"(x) \sim \frac{f(x) - hf'(x) + \frac{h^2}{2!}f''(x) - \frac{h^3}{3!}f'''(x) + \frac{h^4}{4!}f''''(\xi) + f(x) + hf'(x) + \frac{h^2}{2!}f''(x) + \frac{h^3}{3!}f'''(x) + \frac{h^4}{4!}f''''(\xi) - 2f(x)}{h^2}$$

$$f"(x) \sim f"(x) + \frac{h^2}{12}f''''(\xi)$$

So, algorithm error $= \frac{h^2}{12}f''''(\xi)$

Machine error = round off error There are round off error in f(x-h), f(x+h) and f(x), thus we expect round off error be

$$\approx \frac{\varepsilon f(x-h) + \varepsilon f(x+h) + \varepsilon 2f(x)}{h^2} \approx \frac{\sqrt{1+1+2^2}\,\varepsilon f(x)}{h^2} = \frac{\sqrt{6}\,\varepsilon f(x)}{h^2}$$

Total error

$$= \frac{h^2}{12}f''''(\xi) + \frac{\sqrt{6}\,\varepsilon f(x)}{h^2}$$

b. The optimum value of h occurs when the two error sources are of the same magnitude:

$$\frac{\sqrt{6}\,\varepsilon f(x)}{h_{opt}^2} \approx \frac{h_{opt}^2}{12}f''''(\xi)$$

$$h_{opt} \approx \left[\frac{12\sqrt{6}\,\varepsilon f(x)}{f''''(\xi)}\right]^{\frac{1}{4}}$$

Total minumum possible error

$$= \frac{h_{opt}^2}{12}f''''(\xi) + \frac{\sqrt{6}\,\varepsilon f(x)}{h_{opt}^2}$$

$$= \frac{1}{3}\left[3\sqrt{6}\,\varepsilon f(x) f''''(\xi)\right]^{\frac{1}{2}}$$

Problem 4. Gaussian Elimination:

Given the 2x2 matrix A shown below, compute the upper bound on the error of the solution x of Ax=b given that we have some error in the right hand side vector b.

A = [1, 2; 3, 7]

Hint: the formula for inv(A) is given by:

A = [a, b; c, d]   inv(A) = [d, -b; -c, a] / (ad - bc)

Solution

To see how much x changes with a small change in b

$$Ax = b$$

$$A(x + \Delta x) = (b + \Delta b)$$

$$\Longrightarrow \quad A\Delta x = \Delta b$$

$$\Delta x = A^{-1}\Delta b$$

$$||\Delta x|| = ||A^{-1}\Delta b|| \leq ||A^{-1}||\ ||\Delta b||$$

$$b = Ax \quad \Rightarrow \quad ||b|| = ||Ax||$$

$$||b|| \leq ||A||\ ||x||$$

$$\frac{1}{||b||} \geq \frac{1}{||A||\ ||x||}$$

the relative error in x is measured by the ratio $||\Delta x||/||x||$

$$\left.\begin{array}{rcl} ||\Delta x|| & \leq & ||A^{-1}||\ ||\Delta b|| \\ \\ \dfrac{1}{||b||} & \geq & \dfrac{1}{||A||\ ||x||} \end{array}\right\} \quad \times$$

$$\frac{||\Delta x||}{||A||\ ||x||} \leq \frac{||A^{-1}||\ ||\Delta b||}{||b||}$$

$$\frac{||\Delta x||}{||x||} \leq ||A||\ ||A^{-1}||\frac{\Delta b||}{||b||}$$

$$\frac{||\Delta x||}{||x||} \leq cond(A)\frac{||\Delta b||}{||b||}$$

Manhattan norm,

$$\|A\|_1 = \max_j \sum_{i=1}^{m} |a_{ij}| \qquad (maximum\ absolute\ column\ sum)$$

$$A = \begin{vmatrix} 1 & 2 \\ 3 & 7 \end{vmatrix} \quad \Longrightarrow \|A\| = 9$$

$$A^{-1} = \begin{vmatrix} 7 & -2 \\ -3 & 1 \end{vmatrix} \quad \Longrightarrow \|A^{-1}\| = 10$$

$$cond(A) = \|A\|\ \|A^{-1}\| = 90$$

$$\frac{\|\Delta x\|}{\|x\|} \leq 90 \frac{\|\Delta b\|}{\|b\|}$$

Problem 5. Statistics and Probability:

Suppose we have a set of N measurements [x(1), x(2), ... x(N)] for which we wish to calculate the variance. Prove that an unbiased estimate of the population variance is given by:

$$\sigma^2 = (\sum_{i=1}^{N} {x_i}^2 - \frac{\sum_{i=1}^{N} {x_i}^2}{N})/(N-1)$$

Justify the N-1 factor.

Solution:

Sample variance is defined as:

$${S_x}^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \mu)^2$$

where $\chi = (\sum_{i=1}^{N} x_i)/\mathrm{N}$

$$= \frac{1}{N-1} \sum_{i=1}^{N} ({x_i}^2 - 2x_i\chi + \chi^2)$$

$$= \frac{1}{N-1} (\sum_{i=1}^{N} {x_i}^2 - 2\chi \sum_{i=1}^{N} x_i + \sum_{i=1}^{N} \chi^2)$$

$$= \frac{1}{N-1}(\sum_{i=1}^{N} {x_i}^2 - 2N\chi\chi + N\chi^2)$$

$$= \frac{1}{N-1}(\sum_{i=1}^{N} {x_i}^2 - N\chi^2)$$

$$= \frac{1}{N-1}\sum_{i=1}^{N}({x_i}^2 - \chi^2)$$

$$= \frac{1}{N-1}\sum_{i=1}^{N}({x_i}^2 - (\sum_{i=1}^{N} x_i)^2/N)^2)$$

$$= \frac{1}{N-1}(\sum_{i=1}^{N} {x_i}^2 - (\sum_{i=1}^{N} x_i)^2/N)$$

Now,

$$\sum_{i=1}^{N} (x_i - \mu)^2 = \sum_{i=1}^{N} (x_i - \chi + \chi - \mu)^2$$

$$= \sum_{i=1}^{N} (x_i - \chi)^2 - 2(\chi - \mu)\sum_{i=1}^{N}(x_i - \chi) + \sum_{i=1}^{N} (\chi - \mu)^2$$

By definition, $\sum_{i=1}^{N}(x_i - \chi) = 0$

Also we have,

$$E((\chi - \mu)^2) = \sigma^2/N$$

and

$$E((x_i - \mu)^2) = \sigma^2$$

$$E({S_x}^2) = E(\frac{1}{N-1}\sum_{i=1}^{N} (x_i - \chi)^2)$$

$$= \frac{1}{N-1}(E(\sum_{i=1}^{N} (x_i - \mu)^2) - NE((\chi - \mu)^2))$$

$$= \frac{1}{N-1}(N\sigma^2 - \frac{N}{N}\sigma^2) = \sigma^2$$

So ${S_x}^2$ is an unbiased estimate for σ^2.

The N-1 part comes about because we are calculating the mean from the same sample data that had been used in calculating the variance. We have reduced the number of degrees of freedom by one.