Estimation

John Lee

UCLA NanoCAD

June 19, 2011

Introduction

What is estimation?

> *"Estimation is the process of extracting information from data - data which can be used to infer the desired information and may contain errors."*
> *– Arthur Gelb, Applied Optimal Estimation*

Idea: extract information from noisy data

Linear least squares and least norm estimates

- Assume no information on the probability distribution of the error
- Work with linear equations:

$$Ax = b \tag{1}$$

  where A and b are known, and x is a vector of unknowns
- Examples
  - Linear least-squares fit of a data – fit a line to data
  - Minimum norm fit – find the parameters that fit the data with minimum norm

Linear least squares

Suppose we have an overdetermined set of equations:

$$Ax = b \tag{2}$$

where $|b| > |x|$ (there are more data points b than parameters x). The estimate $\hat{x}$ that minimizes $||Ax - b||_2$ is given by:

$$\hat{x} = (A^T A)^{-1} A^T b \tag{3}$$

Proof.

$$||Ax - b||_2 = x^T A^T Ax - 2b^T Ax + b^T b \tag{4}$$

Taking the derivative gives:

$$2A^T Ax - 2A^T b = 0 \tag{5}$$

which gives us the desired result □

Minimum norm

Suppose we have an *under*determined set of equations:

$$Ax = b \tag{6}$$

where $|b| < |x|$ (there are less points b than parameters x). The idea is to find $\hat{x}$ that satisfies $Ax = b$ and minimizes $||x||_2$. This is given by:

$$\hat{x} = A^T(AA^T)^{-1}b \tag{7}$$

Proof is left as an exercise.

Minimum Variance Estimates

Find the estimate that minimizes the variance of the error

Example

Given two measurements, m_1 and m_2 of m, each with independent, zero mean Gaussian measurement errors, σ_1 and σ_2, find the minimum variance estimate $\hat{m}$ using the form:

$$\hat{m} = \alpha_1 m_1 + \alpha_2 m_2 \tag{8}$$

Example, contd

Solution

1. *Because the solutions are unbiased,* $\alpha_1 + \alpha_2 = 1$ *(otherwise there will be an error in the mean value).*
2. *The variance of* $\hat{m}$ *is*

$$\alpha_1^2\sigma_1^2 + \alpha_2^2\sigma_2^2 = (1-\alpha_2)^2\sigma_1^2 + \alpha_2^2\sigma_2^2 \tag{9}$$

Minimizing this gives $-2(1-\alpha_2)\sigma_1^2 + 2\alpha_2\sigma_2^2 = 0$ *and*

$$\alpha_2 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \tag{10}$$

Gauss-Markov Estimate

In general, suppose that we have vectors y, β and ϵ, and a measurement matrix W related by

$$y = W\beta + \epsilon \tag{11}$$

y is the measured quantity, β are the underlying systen parameters, and ϵ is the measurement noise, with covariance Q.
The minimum variance unbiased estimate is

$$\hat{\beta} = (W^T Q^{-1} W)^{-1} W^T Q^{-1} y \tag{12}$$

with covariance

$$\mathbf{E}[(\hat{\beta} - \beta)(\hat{\beta} - \beta)^T] = (W^T Q^{-1} W)^{-1} \tag{13}$$

General Minimum Variance Estimate

Suppose, as in the previous slide, that

$$y = W\beta + \epsilon \tag{14}$$

y is the measured quantity, β are the underlying systm parameters, and

$$\mathrm{Cov}[\epsilon\epsilon^T] = Q \tag{15}$$

$$\mathrm{Cov}[yy^T] = P \tag{16}$$

e.g, there is errors in y. The minimum variance unbiased estimate is

$$\hat{\beta} = (P^{-1} + W^T Q^{-1} W)^{-1} W^T Q^{-1} y \tag{17}$$

Maximum likelihood estimates

Another approach to finding estimates – choose estimate $\hat{x}$ that is the most likely

Steps:

1. 1. Write the probability distribution function for x.
2. 2. Maximize the probability distribution function
3. $\rightarrow$ the Maximizer is the ML estimate

Example

Given two measurements, m_1 and m_2 of m, each with independent Gaussian measurement errors, σ_1 and σ_2, find the maximum likelihood estimate $\hat{m}$.

Solution

The pdf of m_1 and m_2 is given by:

$$p(m) = \prod_i \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma_i^2}(m_i - m)^2\right) \tag{18}$$

$$\tag{19}$$

Taking the log yields:

$$p(m) = -\sum_i (\log(\sigma_i \sqrt{2\pi})) + \sum_i \left(-\frac{1}{2\sigma_i^2}(m_i - m)^2\right) \tag{20}$$

$$\tag{21}$$

Example contd

The ML estimate is the value of m that maximizes:

$$p(m) = -\sum_i \log(\sigma_i \sqrt{2\pi}) + \sum_i \left(-\frac{1}{2\sigma_i^2}(m_i - m)^2 \right) \tag{22}$$

Differentiating with respect to m gives:

$$\frac{1}{\sigma_1^2}(m_1 - m) + \frac{1}{\sigma_2^2}(m_2 - m) = 0 \tag{23}$$

and that

$$\hat{m} = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} m_1 + \frac{\sigma_2^1}{\sigma_1^2 + \sigma_2^2} m_2 \tag{24}$$

This is equivalent to the minimum variance estimate.