

Stat 471: Lecture 9

Least Squares Estimation.

Moo K. Chung
mchung@stat.wisc.edu

September 24, 2003

1.

$$\begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The least squares solution is given by

```
>> X=[2 1
      4 2];
>> pinv(X)*[1 1]'
ans =
    0.2400
    0.1200
```

2. Response function estimation. For given bivariate data (x_i, y_i) . Let's see if the following model fits data.

$$Y_i = \phi(x_i) + \epsilon_i,$$

where ϵ_i is i.i.d. Gaussian noise. A more sophisticated correlated error modeling is also possible (possible project topic). Some possible model choice would be $\phi(x) = \beta_0 + \beta_1 x$ (linear), $\phi(x) = \beta_0 + \beta_1 x + \beta_2 x^2$ (quadratic). In general for basis functions $\phi_i(x)$, we may let $\phi(x) = \sum_{i=0}^m \beta_i \phi_i(x)$ and we estimate β_i via the least squares estimation. We need to solve the system of the linear equation

$$\begin{pmatrix} \phi_1(x_1) & \cdots & \phi_m(x_1) \\ \phi_1(x_2) & \cdots & \phi_m(x_2) \\ \vdots & \vdots & \vdots \\ \phi_1(x_n) & \cdots & \phi_m(x_n) \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

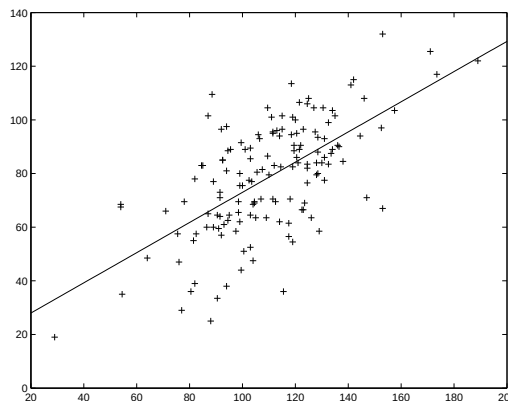
Then we solve the normal equation

$$X'X\beta = X'y$$

If there are more data then the basis functions, i.e. $n \geq m$, $X'X$ is invertible. So we get

$$\hat{\beta} = (X'X)^{-1}X'y.$$

Popular choice for ϕ_i is Hermite polynomials $H_i(x)$ which forms orthogonal polynomial basis



with respect to e^{-x^2} in R , i.e.

$$\int_{-\infty}^{\infty} H_i(x)H_j(x)e^{-x^2} dx = \delta_{ij}2^j j! \sqrt{\pi}.$$

The first few Hermite polynomials are $H_0(x) = 1$, $H_1(x) = 2x$, $H_2(x) = 4x^2 - 2$.

3. Linear regression. Let $\phi(x) = \beta_0 + \beta_1 x$. Download data <http://www.stat.wisc.edu/~mchung/teaching/data/strength.data> and save it into your MATLAB working directory or alternately simply copy and paste into your editor.

```
load strength.data
arm=strength(:,1); grip=strength(:,2)
plot(arm,grip,'+');hold on
X=[ones(147,1) arm]
beta=pinv(X)*y
beta=inv(X'*X)*X'*y
>>beta=X\y
beta =
    16.7297
     0.5627
i=[20:200]
plot(i,beta(1) + beta(2)*i)
```

Chapter 10 nonparametric regression is related to Lecture 9. Next week we will discuss about Monte-Carlo method (Chapter 6). Be prepared!