STA 5106 Computational Methods in Statistics I

Department of Statistics
Florida State University

Class 2
August 29, 2019

Review: Multiple Linear Regression

- The observations may belong to different sample times t_1, $t_2, \ldots, t_m$ such that

$$y(t_i) = \sum_{j=1}^{n} b_j x_j(t_i) + \varepsilon(t_i)$$

- In a matrix form these equations can be restated as

$$y = Xb + \varepsilon$$

where

$$y = \begin{pmatrix} y(t_1) \\ y(t_2) \\ \vdots \\ y(t_m) \end{pmatrix} \quad X = \begin{pmatrix} x_1(t_1) & x_2(t_1) & & x_n(t_1) \\ x_1(t_2) & x_2(t_2) & & x_n(t_2) \\ & & \ddots & \\ x_1(t_m) & x_2(t_m) & & x_n(t_m) \end{pmatrix} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \quad \varepsilon = \begin{pmatrix} \varepsilon(t_1) \\ \varepsilon(t_2) \\ \vdots \\ \varepsilon(t_m) \end{pmatrix}$$

Review: Least Squares Solution

- **Least Squares:** The goal is to find the weight vector b such that

$$\| y - Xb \|^2$$

is minimized.

- **This is equivalent to Maximum Likelihood Estimation** of b with model $y = Xb + \varepsilon$, where $\varepsilon \sim N(0, \sigma^2 I)$:

- In order to make the system identifiable, we often have $m > n$.

- In case $X^T X$ is non-singular, the solution is

$$\hat{b} = (X^T X)^{-1} X^T y$$

The term $(X^T X)^{-1} X^T$ is called the **pseudo-inverse** of X.

Review: Orthogonal Transformations

- Let Q be an $m \times m$ orthogonal matrix, i.e. $QQ^T = Q^TQ = I_m$, and

$$y^* = Qy = QXb + Q\varepsilon = X^*b + \varepsilon^*$$

- If Q is chosen such that X^* is an **upper triangular** matrix, then

$$\begin{pmatrix} y_1^* \\ y_2^* \end{pmatrix} = \begin{pmatrix} X_1^* \\ 0 \end{pmatrix} b + \begin{pmatrix} \varepsilon_1^* \\ \varepsilon_2^* \end{pmatrix}$$

- Therefore,

$$\| y^* - X^* b \|^2 = \| y_1^* - X_1^* b \|^2 + \| y_2^* \|^2$$

and

$$\hat{b} = \arg\min_b \| y - Xb \|^2 = \arg\min_b \| y^* - X^* b \|^2$$

$$= \arg\min_b \| y_1^* - X_1^* b \|^2$$

Backward Substitution

- This is a technique to solve for the vector *b* in the linear equation:

$$\begin{pmatrix} y_1^* \\ y_2^* \\ \vdots \\ y_n^* \end{pmatrix} = \begin{pmatrix} x_{11}^* & x_{12}^* & & x_{1n}^* \\ 0 & x_{22}^* & & x_{2n}^* \\ & & \ddots & \\ 0 & 0 & & x_{nn}^* \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

- At first,

$$\hat{b}_n = y_n^* / x_{nn}^*$$

- Then,

$$\hat{b}_{n-1} = (y_{n-1}^* - x_{n-1,n}^* \hat{b}_n) / x_{n-1,n-1}^*$$

- In general,

$$\hat{b}_j = (y_j^* - \sum_{i=j+1}^{n} x_{j,i}^* \hat{b}_i) / x_{j,j}^*, \qquad j = n-1, n-2, \ldots, 1$$

Algorithm

- Given an upper triangular matrix $X \in \mathbf{R}^{n\times n}$ and a vector $y \in \mathbf{R}^{m}$, find the least squares estimate of b.

- **Algorithm 8 (Backward Substitution)**

```
function b = backsub(X,y)
l = size(X);
n = l(2);
b(n, 1) = y(n, 1)/X(n, n);
for j = n - 1 : -1 : 1
        b(j, 1) = (y(j, 1) - X(j, j + 1 : n) * b(j + 1 : n, 1))/X(j, j);
end
```

Upper Triangular Matrix

- How to find O efficiently such that X^* is upper triangular?
- Two popular techniques to find such orthogonal transformations: **Householder transformations** and Givens rotations.
- These are based on modifying the column vectors of matrix X using the reflections or the rotations.
- By selection of appropriate plane of reflections or appropriate angles of rotation, we can introduce zeros in appropriate places in X.

Basic Rotations and Reflections

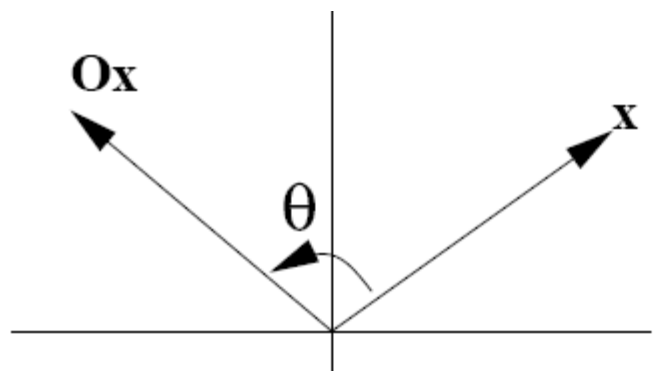
- In our case, for a $m \times n$ matrix X (with $m > n$) we want to zero out the elements below the diagonal elements.
- **Definition 6** An $n \times n$ matrix O is called an orthogonal matrix if $O^T O = I_n$, where I_n is an $n \times n$ identity matrix.
- There are two types of orthogonal matrices: (i) rotations and (ii) reflections.
- Since for an orthogonal matrix O, $O^T O = I_n$, we have $\det(O) = \pm 1$.
- In case the **determinant is +1**, the matrix O is called a **rotation** matrix, and for **determinant −1**, it is called a **reflection** matrix.

Rotation when $n = 2$

- In case of $n = 2$, the rotation matrices always take the form,

$$O_{rot} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

- For an $x \in \mathbf{R}^2$, the vector Ox is nothing but a rotated version of x, rotated in the counterclock direction by an angle θ.



$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \cos(a) \\ \sin(a) \end{pmatrix} = \begin{pmatrix} \cos(\theta + a) \\ \sin(\theta + a) \end{pmatrix}$$

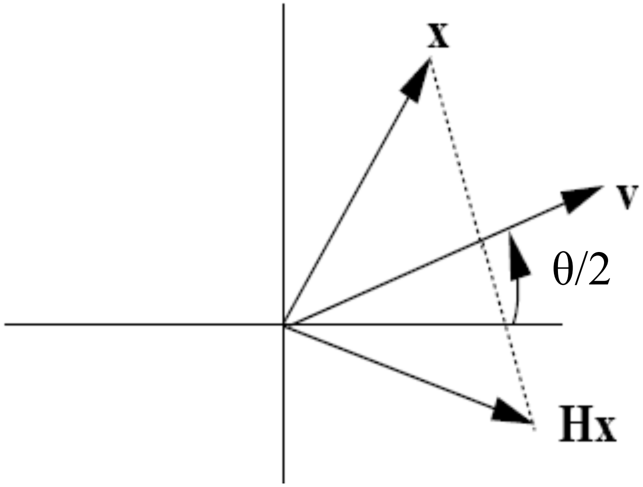
Reflection when $n = 2$

- Instead, the reflection matrices always take the form,

$$H_{ref} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{bmatrix}$$

- Multiplication of a vector by this matrix results in reflection of this vector in a mirror positioned on line $\theta/2$.

$$\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix} \begin{pmatrix} \cos(a) \\ \sin(a) \end{pmatrix} = \begin{pmatrix} \cos(\theta - a) \\ \sin(\theta - a) \end{pmatrix}$$



2.3 Householder Reflection Transformations

Householder Transformation

- **Definition 7** For a vector $v \in \mathbf{R}^m$, an $m \times m$ matrix H of the form

$$H = I_m - 2vv^T/(v^T v)$$

is called a **Householder reflection matrix** (or just Householder matrix or Householder transformation).

- v is called the Householder vector and H is an orthogonal matrix.

- A simple calculation shows that $Hv = -v$.

- It should be noted that the scale (length) of v does not affect H.