LINEAR ALGEBRA

UE19MA251

APARNA B S
Department of Science and Humanities

LINEAR ALGEBRA – UE19MA251

Agenda

- Least square problem – Normal equations.
- Positive Definiteness and the Least squares
- A simple example

LINEAR ALGEBRA – UE19MA251

The Normal Equations

Theorem: *Any solution x of the least squares problem is a solution of the linear system*

$$A^T A x = A^T b.$$

The system is nonsingular if and only if A has linearly independent columns.

Proof.

- Since $b - Ax \in \ker(A^T)$, we have $A^T(b - Ax) = 0$ or $A^T A x = A^T b$.
- $A^T A$ is nonsingular. Suppose $A^T A x = 0$ for some $x \in \mathbb{R}^n$. Then $0 = x^T A^T A x = (Ax)^T A x = \|Ax\|_2^2$. Hence $Ax = 0$ which implies that $x = 0$ if and only if A has linearly independent columns.

The linear system $A^T A x = A^T b$ is called the **normal equations.**

We have learned that **least square** comes from **projection** :

$$b - p = e \Rightarrow, A^T(b - A\hat{x}) = 0 \Rightarrow, A^T A\hat{x} = A^T b$$

Consequently, only if $A^T A$ is invertible, then we can use linear regression to find approximate solutions $\hat{x} = (A^T A)^{-1} A^T b$ to unsolvable systems of linear equations.

According to the reasoning before, we know as long as all columns of $A_{m\times,n}$ are mutual independent, then $A^T A$ is invertible. At the same time we ought to notice that the columns of A are guaranteed to be independent if they are orthoganal and even orthonormal.

In another prospective, if $A^T A$ is positive definite, then $A_{m\times,n}$ has rank n (independent columns) and thus $A^T A$ is invertible.

Overall, if $A^T A$ is positive definite or invertible, then we can find approximate solutions of least square.

Find the function of the form

$$F(x) = c_1 + c_2 x \lg x + c_3 e^x$$

that is the best least-squares fit to the data points

$$(1,1), (2,1), (3,3), (4,8).$$

First we form the A matrix

$$A = \begin{pmatrix} 1 & 0 & e \\ 1 & 2 & e^2 \\ 1 & 3\lg 3 & e^3 \\ 1 & 8 & e^4 \end{pmatrix}$$

We compute the pseudoinverse, then multiply it by y, to obtain the coefficient vector

$$c = \begin{pmatrix} 0.411741 \\ -0.20487 \\ 0.16546 \end{pmatrix}.$$

- Given $A^{m,n}$ and $b \in \mathbb{R}^m$.
- The system $Ax = b$ is over-determined if $m > n$.
- This system has a solution if $b \in \text{span}(A)$, the column space of A, but normally this is not the case and we can only find an approximate solution.
- A general approach is to choose a vector norm $\|\cdot\|$ and find x which minimizes $\|Ax - b\|$.
- We will only consider the Euclidian norm here.

- Given $A^{m,n}$ and $b \in \mathbb{R}^m$ with $m \geq n \geq 1$. The problem to find $x \in \mathbb{R}^n$ that minimizes $\|Ax - b\|_2$ is called the least squares problem.
- A minimizing vector x is called a least squares solution of $Ax = b$.
- Several ways to analyze:
- Quadratic minimization
- Orthogonal Projections
- SVD

LINEAR ALGEBRA – UE19MA251

Positive Definite Matrices and Least Squares

- Define function $E : \mathbb{R}^n \to \mathbb{R}$ by $E(\boldsymbol{x}) = \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2$
- $E(\boldsymbol{x}) = (\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) = \boldsymbol{x}^T\boldsymbol{B}\boldsymbol{x} - 2\boldsymbol{c}^T\boldsymbol{x} + \alpha$, where
- $\boldsymbol{B} := \boldsymbol{A}^T\boldsymbol{A}$, $\boldsymbol{c} := \boldsymbol{A}^T\boldsymbol{b}$ and $\alpha := \boldsymbol{b}^T\boldsymbol{b}$.
- $\boldsymbol{B}$ is positive semidefinite and positive definite if $\boldsymbol{A}$ has rank n.
- Since the Hessian $\boldsymbol{H}E(\boldsymbol{x}) := \left(\frac{\partial^2 E(\boldsymbol{x})}{\partial x_i \partial x_j}\right) = 2\boldsymbol{B}$ we can find minimum by setting partial derivatives equal zero.
- $\nabla E(\boldsymbol{x}) := \left(\frac{\partial E(\boldsymbol{x})}{\partial x_i}\right) = 2(\boldsymbol{B}\boldsymbol{x} - \boldsymbol{c}) = \boldsymbol{0}$
- Normal equations $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$.

LINEAR ALGEBRA – UE19MA251

A simple example

- $$x_1 = 1$$
  $$x_1 = 1, \quad \boldsymbol{A} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{x} = [x_1], \quad \boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix},$$
  $$x_1 = 2$$
- Quadratic minimization problem:
  $$\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 = (x_1 - 1)^2 + (x_1 - 1)^2 + (x_1 - 2)^2.$$
- Setting the first derivative with respect to x_1 equal to zero we obtain $2(x_1 - 1) + 2(x_1 - 1) + 2(x_1 - 2) = 0$ or $6x_1 - 8 = 0$ or $x_1 = 4/3$
- The second derivative is positive (it is equal to 6) and $x = 4/3$ is a global minimum.

THANK YOU

Aparna B. S
Department of Science & Humanities
aparnabs@pes.edu