Lecture 19

Fundamental Theorem of Linear Algebra & Least-Square Approximation

Fundamental Subspaces

Let $A \in M_{m\times n}(\mathbb{R})$. Suppose $N(A)$ is the null space of A, $C(A)$ is the column space of A, $C(A^T)$ is the column space of A^T and $N(A^T)$ is the null space of A^T. Then $N(A)$, $C(A^T)$ are subspaces of $\mathbb{R}^n$, and $C(A)$, $N(A^T)$ are subspaces of $\mathbb{R}^m$. These subspaces are called fundamental subspaces associated to A.

Lemma 1. *$N(A) \perp C(A^T)$ and $C(A) \perp N(A^T)$.*

Proof: Let $x \in N(A)$ and $y \in C(A^T)$. Then $A(x) = 0$ and $A^T z = y$ for some $z \in \mathbb{R}^m$. Then $y^T x = z^T A x = 0$, that is, $\langle x, y\rangle = 0$ so that $N(A) \perp C(A^T)$. Similarly, $C(A) \perp N(A^T)$. □

Theorem 2 (Fundamental Theorem of Linear Algebra)**.** *Let $A \in M_{m\times n}(\mathbb{R})$. Then*
1. $\mathbb{R}^n = N(A) \oplus C(A^T)$
2. $\mathbb{R}^m = C(A) \oplus N(A^T)$.

Proof: Since $C(A^T)$ is a subspace of $\mathbb{R}^n$, $\mathbb{R}^n = C(A^T) \oplus (C(A^T))^{\perp}$. We claim that $C(A^T)^{\perp} = N(A)$. By Lemma 11, $N(A) \subseteq C(A^T)^{\perp}$. Note that $n = \dim(C(A^T)) + \dim((C(A^T))^{\perp})$ and by rank-nullity theorem $n = rank(A) + nullity(A)$. This implies $\dim(N(A)) = dim((C(A^T))^{\perp})$. Hence, $N(A) = (C(A^T))^{\perp}$. Similarly one can proof $\mathbb{R}^m = C(A) + N(A^T)$. □

Least-Square Approximation

Problem 3. *Let $A \in M_{m\times n}(\mathbb{R})$ and $b \in \mathbb{R}^n$ such that $b \notin C(A)$, where $C(A)$ is the column-space of A. In other words, the system $Ax = b$ is inconsistent. So the problem is to find a "pseudo solution" or "approximate solution" under certain condition in error term.*

Definition 4 (Least-Square Method)**.** *A method to approximate a solution of an inconsistent system of linear equations such that the solution minimizes the sum of square of errors made in every equation.*

Let $AX = b$ be an inconsistent system of linear equation, where $A \in M_{m\times n}(\mathbb{R}), X \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$. Suppose $X_0 = (x_1, x_2, \ldots, x_n)$ is an approximate solution of the system. Then $AX_0 = b'$ and $b' \neq b$. The error term for the i-th equation is $|b_i - b_i'| = |\sum_{j=1}^{n} a_{ij}x_j - b_i|$. For X_0 to be a least-square solution of the system, the sum of square of the errors made in each equation should be minimum, that is,

$$\sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij}x_j - b_i \right|^2 \text{ is minimum.}$$

Theorem 5. *Suppose X_0 is a least square solution. Then AX_0 is the orthogonal projection of b on the column-space of A.*

Proof. Let X_0 be the least-square approximation of $AX = b$. Then $\sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij}x_j - b_i \right|^2$ is minimum. For $Y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$, $\sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij}y_j - b_i \right|^2 = ||AY - b||^2$. Thus $||AX_0 - b|| \leq ||AX - b||$ for all $X \in \mathbb{R}^n$

as $||AX_0 - b||$ is minimum. Recall that w_v is the orthogonal projection of v on to W if and only if $||v - w_v|| \leq ||v - w||$ for all $w \in W$. Take $V = \mathbb{R}^m$, $W = \{AX : X \in \mathbb{R}^n\}$ and $v = b$. Then AX_0 is the orthogonal projection of b on the column-space of A. □

Theorem 6. *Let X_0 be a least-square approximation of $AX = b$ and $N(A)$ be the null space of A. Suppose S is the set of all least-square solutions of $AX = b$. Then $S = X_0 + N(A)$.*

Proof. Let $X \in X_0 + N(A)$. Then $X = X_0 + X_h$ so that $AX - b = AX_0 - b$. Thus $X \in S$. Now suppose $X \in S$. Then $||AX - b|| = ||AX_0 - b|| \Rightarrow ||(AX_0 - b) + A(X - X_0)||^2 = ||AX_0 - b||^2 \Rightarrow ||AX_0 - b||^2 + ||A(X - X_0)||^2 = ||AX_0 - b||$ since $A(X - X_0) \in C(A)$ and $(AX_0 - b) \perp Y$ for all $Y \in C(A)$. Therefore, $||A(X - X_0)|| = 0 \Rightarrow A(X - X_0) = 0 \Rightarrow X - X_0 \in N(A)$ so that $X = X_0 + (X - X_0) \in X_0 + N(A)$. □

Application of Fundamental Theorem of Linear Algebra

Lemma 7. Let $A \in M_{m\times n}(\mathbb{R})$. Then the $A^T AX = A^T b$ is consistent for every $b \in \mathbb{R}^m$.

Proof. It is enough to show that each $A^T b$ is in the column space of $A^T A$. By Fundamental Theorem of Linear Algebra, $\mathbb{R}^m = C(A) \oplus N(A^T)$. Thus, there exist $X \in \mathbb{R}^m$ and $Y \in N(A^T)$ such that $b = AX + Y$. Therefore, $A^T b = A^T(AX) + A^T Y = A^T AX + 0$. □

Theorem 8. *Let $AX = b$ be an inconsistent system of linear equations and $X_0 \in \mathbb{R}^n$. Then X_0 is a least-square solution of $AX = b$ if and only if $A^T AX_0 = A^T b$.*

Proof. Note that $N(A^T)^\perp = C(A)$. Then X_0 is a least-square solution if and only if $AX_0 - b \in C(A)^\perp$, that is, $(AX_0 - b) \in N(A^T) \Leftrightarrow A^T(AX_0 - b) = 0 \Leftrightarrow A^T AX_0 = A^T b$.

□

Remark 9. *For finding a least-square solution, one can solve the system $A^T AX = A^T b$.*

Example 10. *Find a straight line $y = a + bx$ which fits best the given points $(1, 0), (2, 3), (3, 4), (4, 4)$ by least-square method.*

Solution: We get the following system of equations

$$\begin{aligned} a + b &= 0 \\ a + 2b &= 3 \\ a + 3b &= 4 \\ a + 4b &= 4 \end{aligned}$$

which is inconsistent. For finding a least-square solution, we will solve the system $A^T AX = A^T b$, where $A \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix}$ and $b = \begin{pmatrix} 0 \\ 3 \\ 4 \\ 4 \end{pmatrix}$. Thus, $A^T A = \begin{pmatrix} 4 & 10 \\ 10 & 30 \end{pmatrix}$ and $A^T b = \begin{pmatrix} 11 \\ 34 \end{pmatrix}$. Thus, $(A^T A|A^T b =$ $\begin{pmatrix} 4 & 10 & | & 11 \\ 10 & 30 & | & 34 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & | & 34/10 \\ 4 & 10 & | & 11 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & | & 34/10 \\ 0 & -2 & | & -13/5 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & | & -1/2 \\ 0 & 1 & | & 13/10 \end{pmatrix}$. Thus, $y = -1/2 + 13/10x$ is a best fit.

For applying orthogonal projection method, $W = C(A) = \{(x+y, x+2y, x+3y, x+4y) \mid x, y \in \mathbb{R}\}$. Basis of W is $\{(1,1,1,1),(1,2,3,4)\}$. An orthogonal basis of W is $\{(1,1,1,1),(-3/2,-1/2,1/2,3/2)\}$, $||(1,1,1,1)||^2 = 4$ and $||(-3/2,-1/2,1/2,3/2)||^2 = 5$. Take $v = b = (0,3,4,4)$. Then $P_W(v) = 11/4(1,1,1,1) + 13/10(-3/2,-1/2,1/2,3/2) = 1/10(8,21,34,47)$. Then a least-square solution can be obtained by solving $AX = P_W(v)$ so that $\begin{pmatrix} 1 & 1 & | & 8/10 \\ 1 & 2 & | & 21/10 \\ 1 & 3 & | & 34/10 \\ 1 & 4 & | & 47/10 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & | & 8/10 \\ 0 & 1 & | & 13/10 \\ 0 & 2 & | & 26/10 \\ 0 & 3 & | & 39/10 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & | & 8/10 \\ 0 & 1 & | & 13/10 \\ 0 & 0 & | & 0 \\ 0 & 0 & | & 0 \end{pmatrix}$.
Thus $(-1/2, 13/10)$ is a least-square solution so that $y = -1/2x + 13/10$ is a best fit.