MA-106 Linear Algebra

M.K. Keshari

Department of Mathematics
Indian Institute of Technology Bombay
Powai, Mumbai - 76

29th January, 2015
D1 - Lecture 11

Recall:

Two vectors $v, w \in \mathbb{R}^n$ are orthogonal if $v^T w = 0$

A set of vectors $\{v_1, v_2, \cdots\} \subseteq \mathbb{R}^n$ is an **orthogonal set** if $v_i \neq 0$ for all i and $v_i^T v_j = 0$ for all $i \neq j$.

Every orthogonal set in $\mathbb{R}^n$ is a linearly independent set.

Subspaces V and W of $\mathbb{R}^n$ are orthogonal, $V \perp W$, if every vector in V is orthogonal to every vector in W.

If $V \perp W$ are orthogonal subspaces of $\mathbb{R}^n$, then
$\dim V + \dim W \leq n$.
Further, we can enlarge W to W', such that
$V \perp W'$ and $\dim V + \dim W' = n$.

Fundamental Theorem of Orthogonality - 1
A : $m \times n$ matrix, then row space of A is orthogonal to $N(A)$, and $C(A)$ is orthogonal to $N(A^T)$.

Orthogonal Complements

Definition. Let W be a subspace of $\mathbb{R}^n$. Its **orthogonal complement** $W^\perp = \{v \in \mathbb{R}^n \mid v^T w = 0 \text{ for all } w \in W\}$.

Ex: $W^\perp$ is a subspace of $\mathbb{R}^n$,
i.e. $v_1, v_2 \in W^\perp$ and $c_1, c_2 \in \mathbb{R} \implies (c_1 v_1 + c_2 v_2) \in W^\perp$.

Theorem (Fundamental Theorem of Orthogonality - 2)

Let A be an $m \times n$ matrix.
1. Orthogonal complement of row space of $A = N(A)$.
2. Orthogonal complement of column space of $A =$ left nullspace $N(A^T)$.

F.T.O.-1: $C(A^T) \perp N(A) \Rightarrow N(A) \subset C(A^T)^\perp$.
F.T.O.-2: $\implies N(A) = C(A^T)^\perp$.

Theorem: (Orthogonal Complement)

$W \subseteq \mathbb{R}^n$: subspace $\Rightarrow$ dim $W +$ dim $W^\perp = n$.

Proof. Let $v_1, \ldots, v_r$ be a basis of W. Let A be a matrix with rows $v_1, \ldots, v_r$. Then $\text{rank}(A) = r$ and W is row space of A. $W^\perp = N(A)$ is of dimension $n - r$. This proves the theorem.

Observation. $W + W^\perp = \mathbb{R}^n$ and $W \cap W \perp = 0$. Hence every $v \in \mathbb{R}^n$ can be uniquely expressed as $v = w + w'$, where $w \in W$ and $w' \in W^\perp$.

Therefore, if B is a basis of W and B' is a basis of $W^\perp$, then $B \cup B'$ is a basis of $\mathbb{R}^n$.

Some Consequences.

Let A be $m \times n$ of $\operatorname{rank}(A) = r$. Then

1. $C(A^T)^{\perp} = N(A) \Rightarrow \dim(C(A^T)) = r$ and $\dim(N(A)) = n - r$.
2. $C(A)^{\perp} = N(A^T) \Rightarrow \dim(C(A)) = r$ and $\dim(N(A^T)) = m - r$.
3. $C(A^T) \cap N(A) = 0$ and $\mathbb{R}^n = C(A^T) + N(A)$.

Similarly $C(A) \cap N(A^T) = 0$ and $\mathbb{R}^m = C(A) + N(A^T)$.

4. Every $x \in \mathbb{R}^n$ can be written uniquely as $x = x_r + x_n$, where $x_r \in C(A^T)$, $x_n \in N(A)$. Hence $Ax = Ax_r \in C(A)$.

$\boxed{\text{The matrix } A \text{ transforms its row space into its column space.}}$

Orthogonal basis are good

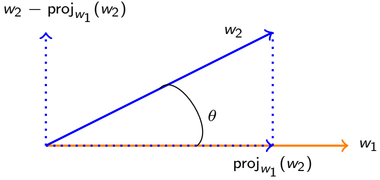
Question. If $v = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ and $w = \begin{pmatrix} 1 \\ 1 \end{pmatrix}\}$, then $\mathcal{B} = \{v, w\}$ is a basis of $\mathbb{R}^2$. How to write $\begin{pmatrix} a & b \end{pmatrix}^T \in \mathbb{R}^2$ as a linear combination of v and w?
For this, we need to solve the system $c_1 v + c_2 w = (a \ \ b)^T$ by Gaussian eliminationq.

If we take an orthogonal basis for $\mathbb{R}^2$, e.g. $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$,
$w = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, then to solve $c_1 v + c_2 w = (a \ \ b)^T$,
$v^T(c_1 v + c_2 w) = v^T(a \ \ b)^T \Rightarrow c_1 = v^T(a \ \ b)/v^T v = (a + b)/2$.
Similarly, $c_2 = w^T(a \ \ b)/w^T w = (-a + b)/2$.

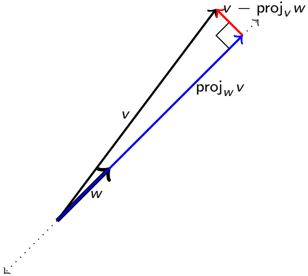
Orthogonal basis: How?

Let W be a subspace of $\mathbb{R}^n$ with basis $B = \{w_1, w_2\}$. It is good to have an orthogonal basis of W.

Question. How to find an orthogonal basis of W?



Question. How to find the orthogonal projection of $v \in \mathbb{R}^n$ onto the line spanned by $w \in \mathbb{R}^n$, denoted by $\text{proj}_w v$.



$\text{proj}_w v = kw$ for some $k \in \mathbb{R}$.

$$\begin{aligned} w^T(v - \text{proj}_w v) &= 0 \\ w^T(v - kw) &= 0 \\ w^T v - kw^T w &= 0 \\ k &= \frac{w^T v}{w^T w} \end{aligned}$$

Thus the orthogonal projection of v on the line Span(w) is

$$\text{proj}_w v = \frac{w^T v}{w^T w}\, w$$

Example. If $b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and $a = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, then find the orthogonal projection of b on the line spanned by a.

The projection is given by

$$\text{proj}_a b = \frac{a^T b}{a^T a} a$$

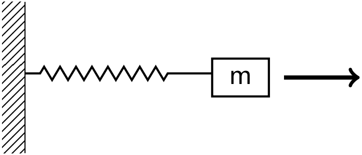
Now $a^T b = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = 6$ and $a^T a = 1 + 4 + 9 = 14$.

Therefore

$$\text{proj}_a b = \frac{6}{14} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

$\square$

Linear Least square with one variable



Hooke's Law states that displacement x of the spring is directly proportional to the load (mass) applied, that is, $m = kx$.

Student A performs experiments to calculate spring constant k. The data collected says for loads $4, 7, 11$ pounds applied, the displacement is $3, 5, 8$ inches respectively. Hence we have the following system.

$$\begin{pmatrix} 3 \\ 5 \\ 8 \end{pmatrix} k = \begin{pmatrix} 4 \\ 7 \\ 11 \end{pmatrix} \qquad (ak = b).$$

Clearly the data is inconsistent. If we allow for various errors, how to give an estimate of the spring constant.

One method is to choose k that minimizes average error E given by $$E^2 = (3k-4)^2 + (5k-7)^2 + (8k-11)^2.$$

For minimum, $dE^2/dx = 0$.

$\Rightarrow 2[(3k-4)3 + (5k-7)5 + (8k-11)8] = 0$

$\Rightarrow (9+25+64)k - (4.3+7.5+11.8) = 0,$

$\Rightarrow a^T ak = a^T b$

$\Rightarrow k = a^T b/a^T a = 135/98.$

The least square solution to the problem $ax = b$ in one unknown is denoted by $\hat{x} = a^T b/a^T a$.

Thus the value of $\hat{k} = 135/98$ minimizes the error.

Linear Least square with several variable

Suppose system $Ax = b$ is inconsistent, i.e. $b \notin C(A)$.
We want to find $\hat{x}$ which minimizes the least square error E.
The error $E = \|Ax - b\|$ is the distance from b to $Ax \in C(A)$.
We want the least square solution $\hat{x}$ which minimizes E. This is same as finding the vector $p = A\hat{x}$ in the column space $C(A)$ which is closest to b.
Thus p must be the orthogonal projection of b onto the column space $C(A)$.
The error vector $e = b - A\hat{x}$ must be perpendicular to $C(A)$.
Since $C(A)^\perp =$ left null space $N(A^T)$,

$$A^T(b - A\hat{x}) = 0 \text{ or } \boxed{A^T A\,\hat{x} = A^T b}$$

Therefore, to find $\hat{x}$, we need to solve $A^T A\,\hat{x} = A^T b$.

Example: Find the least square solution to the system

$$\begin{pmatrix} -1 & 2 \\ 2 & -3 \\ -1 & 3 \end{pmatrix} x = \begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix} \quad (Ax = b)$$

We need to solve $A^T A \hat{x} = A^T b$, i.e.

$$\begin{aligned} \begin{pmatrix} -1 & 2 & -1 \\ 2 & -3 & 3 \end{pmatrix} \begin{pmatrix} -1 & 2 \\ 2 & -3 \\ -1 & 3 \end{pmatrix} \hat{x} &= \begin{pmatrix} -1 & 2 & -1 \\ 2 & -3 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix} \\ \begin{pmatrix} 6 & -11 \\ -11 & 22 \end{pmatrix} \hat{x} &= \begin{pmatrix} -4 \\ 11 \end{pmatrix} \end{aligned}$$

$$\left(\begin{array}{cc|c} 6 & -11 & -4 \\ -11 & 22 & 11 \end{array}\right) \to \left(\begin{array}{cc|c} 6 & -11 & -4 \\ 0 & 22-(121/6) & 11-(44/6) \end{array}\right)$$

$$\begin{aligned} 11/6\, \hat{x}_2 = 11/3 &\implies \hat{x}_2 = 2, \\ 6\, \hat{x}_1 - 22 = -4 &\implies \hat{x}_1 = 3. \end{aligned}$$

$\square$