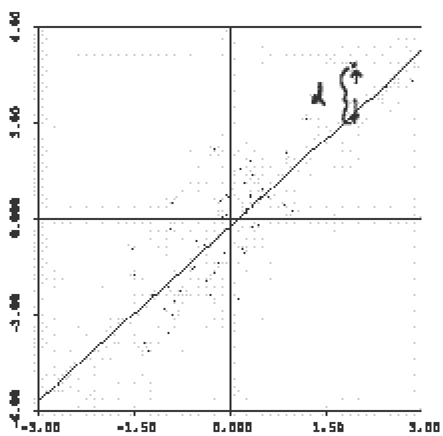
MAT225 Section Summary: 6.1

Inner Product, Length, and Orthogonality
Summary

Our objective here is to solve the least squares problem: there are times when we would like to the equation $A\mathbf{x} = \mathbf{b}$ exactly, but when the solution does not, in fact exist. The question then is, what's the best non-solution? We need to do something, so what should we do when the exact solution isn't a possibility? Do the next best thing....

What do we mean by "next best thing"? We mean that we want to make the distance between $A\mathbf{x}$ and $\mathbf{b}$ as small as possible; that will have to do with definitions of distance, which will fall out of something called an inner product.

The classic example of this is the standard least-squares line, which students of any science are familiar with: In terms of matrix operations, we're



Square d,
sum all the "square ds"
of all the points, &
minimize it.

trying to find coefficients α and β such that

$$\overline{y = \alpha + \beta x}$$

That mini

for all points. Unfortunately, we have more than two points, so the system becomes over-determined:

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

We can't (generally) find an actual solution vector $\mathbf{x} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$ that makes this true, so we make due with an approximate solution $\mathbf{x}^* = \begin{bmatrix} a \\ b \end{bmatrix}$ that gives us a "best fit": that minimizes the distance between the two vectors $A\mathbf{x}^*$ and $\mathbf{y}$.

inner product: The inner product between vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$, or their **dot product**, is defined as

$$\mathbf{u} \cdot \mathbf{v} \equiv \mathbf{u}^T\mathbf{v} = u_1v_1 + u_2v_2 + \ldots + u_nv_n$$

$\bar{u} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ $\quad \bar{u} \cdot \bar{u} = (-1)\cdot(-1) + 2 \cdot 2 = 5$

Example: #1, p. 382.

$\bar{v} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$ $\quad \bar{v} \cdot \bar{v} = 52$ $\quad \dfrac{\bar{v} \cdot \bar{u}}{\bar{u} \cdot \bar{u}} = \dfrac{4\cdot(-1) + 6 \cdot 2}{5} = \dfrac{8}{5}$

Properties of inner products (Theorem 1): Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be vectors in $\mathbb{R}^n$, and c be any scalar. Then

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
2. $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
3. $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (c\mathbf{v})$
4. $\mathbf{u} \cdot \mathbf{u} \geq 0$, and $\mathbf{u} \cdot \mathbf{u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$.

norm: The **length** or **norm** of vector $\mathbf{v}$ is the non-negative scalar $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$. "The shadow it casts on itself."

Example: #7, p. 382.

$\bar{w} = \begin{bmatrix} 3 \\ -1 \\ -5 \end{bmatrix}$ $\quad \|\bar{w}\| = \sqrt{\bar{w} \cdot \bar{w}} = \sqrt{35}$

unit vector: a vector whose length is 1 is called a **unit vector**, and one can "normalize" a vector (that is, give it unit length) by dividing the vector by its norm:

$$\hat{\mathbf{v}} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$$

Example: #9, p. 382.

$\bar{v} = \begin{bmatrix} -30 \\ 40 \end{bmatrix}$; find $\hat{v} = \begin{bmatrix} -30 \\ 40 \end{bmatrix} / 50 = \begin{bmatrix} -3/5 \\ 4/5 \end{bmatrix}$

$\|\hat{v}\| = \sqrt{(-3/5)^2 + (4/5)^2} = \sqrt{\frac{9+16}{25}} = 1$ ✓

---

distance: For **u** and **v** in $\mathbb{R}^n$, the **distance between** u **and** v, denoted dist(u,v), is the length of the vector $\mathbf{u} - \mathbf{v}$. That is,

$$\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

Example: #13, p. 382.

$\bar{x} = \begin{bmatrix} 10 \\ -3 \end{bmatrix}$ $y = \begin{bmatrix} -1 \\ -5 \end{bmatrix}$

$\|\bar{x} - \bar{y}\| = \left\| \begin{bmatrix} 11 \\ 2 \end{bmatrix} \right\| = \sqrt{121+4} = \sqrt{125} = 5\sqrt{5}$

orthogonal: two vectors **u** and **v** in $\mathbb{R}^n$ are **orthogonal** (to each other) if and only if $\mathbf{u} \cdot \mathbf{v} = 0$.

Example: #15, p. 382.

$$\bar{a} = \begin{bmatrix} 8 \\ -5 \end{bmatrix} \qquad \bar{b} = \begin{bmatrix} -2 \\ -3 \end{bmatrix}$$

$$\bar{a} \cdot \bar{b} = -1 \neq 0$$ <u>not orthogonal</u>

Theorem 2 (the Pythagorean Theorem): Two vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

orthogonal complement: If a vector $\mathbf{z}$ is orthogonal to every vector in a subspace W of $\mathbb{R}^n$, then $\mathbf{z}$ is said to be **orthogonal to** W. The set of all such vectors is called the **orthogonal complement of** W, and denoted $W^{\perp}$.

Example: #26, p. 383.

$$\bar{u} = \begin{bmatrix} 5 \\ -4 \\ 7 \end{bmatrix} \qquad W = \{\bar{x} \in \mathbb{R}^3 \,/\, \bar{u} \cdot \bar{x} = 0\}$$
$$= \mathrm{Null}(\bar{u}^T),$$ so it's a subspace.

$$A = [5 \;\; -4 \;\; 7] \qquad A \cdot \bar{x} = [0]$$

Geometrically, the plane through the origin orthogonal to $\bar{u}$.

It is easy to deduce the following facts concerning the orthogonal complement of W:

1. A vector $\mathbf{x}$ is in $W^\perp$ if and only if $\mathbf{x}$ is orthogonal to every vector in a spanning set of W.
2. $W^\perp$ is a subspace of $\mathbb{R}^n$.

Demonstration: #29 and 30, p. 383

Theorem 3: Let A be an m x n matrix. The orthogonal complement of the row space of A is the nullspace of A, and the orthogonal complement of the column space of A is the nullspace of A^T:

$$(\text{Row } A)^\perp = \text{Nul } A \text{ and } (\text{Col } A)^\perp = \text{Nul } A^T$$

The angle between two vectors in $\mathbb{R}^n$ can be defined using the familiar formula from calculus:

$$\cos(\theta) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}$$

One interpretation of the cosine of this angle in higher dimensional space is as a correlation coefficient.

Example:

1. If one vector is defined as the difference between heights of individuals and the mean height, and
2. the other as the difference between weights of individuals and the mean weight, then

---

3. the correlation between the two variables is given by the cosine of this angle.

The correlation coefficient in the case of the figure shown in the linear regression above is .86: the variables are strongly positively correlated.

#29 $W = \text{Span}\{\bar{v}_1, \ldots, \bar{v}_p\}$. Given $\bar{x}$ / $\bar{x} \cdot \bar{v}_i = 0 \ \forall i$. Show that $\bar{x} \in W^\perp$.

Every vector in W can be written as a linear combination of $\bar{v}_1, \ldots, \bar{v}_p$; consider

$\bar{y} = \alpha_1 \bar{v}_1 + \cdots + \alpha_p \bar{v}_p$

$\bar{x} \cdot \bar{y} = \alpha_1 \bar{x} \cdot \bar{v}_1 + \cdots + \alpha_p \bar{x} \cdot \bar{v}_p = 0 + \cdots + 0 = 0$ ✓

So $\bar{x}$ is in $W^\perp$.

#30 Show that $W^\perp$ is a subspace of $\mathbb{R}^n$.

Three things to show: closure under scalar multiplication + addition, + that $\bar{0} \in W^\perp$.

a) Take $\bar{z} \in W^\perp$, $\bar{u} \in W$. So $\bar{z} \cdot \bar{u} = 0$.

Consider $c \cdot \bar{z}$; $(c\bar{z}) \cdot \bar{u} = c(\bar{z} \cdot \bar{u}) = c \cdot 0 = 0$

b) Take $z_1, z_2 \in W^\perp$, $\bar{u} \in W$. So $\bar{z}_i \cdot \bar{u} = 0$.

Then $(z_1 + z_2) \cdot \bar{u} = \bar{z}_1 \cdot \bar{u} + \bar{z}_2 \cdot \bar{u} = 0 + 0 = 0$

c) $\bar{0} \cdot \bar{u} = 0$, so $\bar{0} \in W^\perp$.

Hence $W^\perp$ is a subspace!