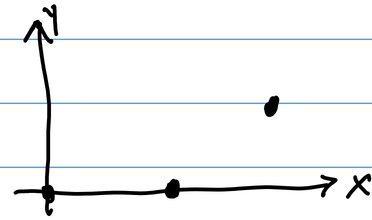
Inner products

We have seen how some systems are inconsistent, but many times in the Real World we are OK with approximate solutions, especially if not accepting one means no solution at all. A least-squares solution $\vec{x}$ to $A\vec{x}=\vec{b}$ is one where $A\vec{x}$ is as close to $\vec{b}$ as possible. "Closeness" is what we aim to define with inner products and orthogonality.

As a preview, to solve the inconsistent system $A\vec{x}=\vec{b}$ approximately, we may solve $A^TA\vec{x} = A^T\vec{x}$ instead.

ex We wish to fit a line $y=mx+b$ to the data

x	y
0	0
1	0
2	1



$$\begin{cases} 0 = m\cdot 0 + b \\ 0 = m\cdot 1 + b \\ 1 = m\cdot 2 + b \end{cases} \longrightarrow \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 3 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 5 & 3 & 2 \\ 3 & 3 & 1 \end{array}\right] \sim \left[\begin{array}{cc|c} 2 & 0 & 1 \\ 1 & 1 & 1/3 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 1/2 \\ 0 & 1 & -1/6 \end{array}\right]$$

$$y = \frac{1}{2}x - \frac{1}{6}$$

Recall now that a vector $\vec{v}$ in $\mathbb{R}^n$ is an $n\times 1$ matrix, so $\vec{v}^T$ is $1\times n$, hence $\vec{v}^T\vec{u}$ is 1×1 ($\vec{u}\in\mathbb{R}^n$, too). Let's agree that 1×1 matrices are just elements of $\mathbb{R}$.

The dot product $\vec{u}\cdot\vec{v}$ of $\vec{u},\vec{v}\in\mathbb{R}^n$ is $\vec{u}^T\vec{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$.

ex
$$\begin{bmatrix}1\\2\\1\end{bmatrix}\cdot\begin{bmatrix}1\\-1\\-2\end{bmatrix} = 1\cdot 1 + 2(-1) + 1(-2) = -3$$

This is an example of an inner product, making $\mathbb{R}^n$ an inner product space. ($C([0,1])$ can have $\langle f,g\rangle = \int_0^1 f(t)g(t)\,dt$; the concept is more general than dot products.)

properties: If $\vec{u},\vec{v},\vec{w}\in\mathbb{R}^n$, $c\in\mathbb{R}$,

(1) $\vec{u}\cdot\vec{v} = \vec{v}\cdot\vec{u}$

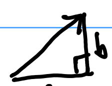
(2) $(\vec{u}+\vec{v})\cdot\vec{w} = \vec{u}\cdot\vec{w} + \vec{v}\cdot\vec{w}$ $\qquad \left((\vec{u}^T+\vec{v}^T)\vec{w} = \vec{u}^T\vec{w} + \vec{v}^T\vec{w}\right)$

(3) $(c\vec{u})\cdot\vec{v} = c(\vec{u}\cdot\vec{v})$

$$\left.\begin{array}{l}(4)\ \ \vec{u}\cdot\vec{u}\geq 0\\(5)\ \ \vec{u}\cdot\vec{u} = 0 \text{ if and only if } \vec{u}=\vec{0}\end{array}\right\}\ \text{positive definite}$$

As for linear combinations,
$$(c_1\vec{u}_1 + \cdots + c_k\vec{u}_k)\cdot\vec{v} = c_1(\vec{u}_1\cdot\vec{v}) + \cdots + c_k(\vec{u}_k\cdot\vec{v})$$

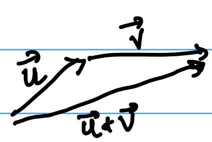
The norm or length of $\vec{v}\in\mathbb{R}^n$ is $\|\vec{v}\| = \sqrt{\vec{v}\cdot\vec{v}} = \sqrt{v_1^2+\cdots+v_n^2}$.
$\vec{v}\cdot\vec{v}$ is "squared length" $\|\vec{v}\|^2$.

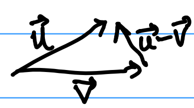
ex  hypotenuse length is $\left\|\begin{bmatrix}a\\b\end{bmatrix}\right\| = \sqrt{a^2+b^2}$.

properties If $\vec{u},\vec{v}\in\mathbb{R}^n$, $c\in\mathbb{R}$,

(1) $\|c\vec{u}\| = |c|\,\|\vec{u}\|$ $\quad (= c\|\vec{u}\| \text{ if } c\geq 0)$

(2) $\|\vec{u}+\vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$ $\quad$ ("triangle inequality")



(3) $\|\vec{u}-\vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$  (again a "triangle inequality")

" $\|\vec{u}+(-\vec{v})\| \leq \|\vec{u}\| + \|-\vec{v}\| = \|\vec{u}\| + \|\vec{v}\|$.

ex if $\|\vec{v}\| < \frac{1}{2}$ and $\|\vec{w}\| < \frac{1}{2}$ then $\|\vec{v}+\vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\| < \frac{1}{2}+\frac{1}{2} = 1$

$(\|\vec{v}+\vec{w}\| < 1)$

A unit vector $\vec{v} \in \mathbb{R}^n$ is a vector with $\|\vec{v}\| = 1$ (unit-length).
The normalization of $\vec{v} \in \mathbb{R}^n$ is $\frac{1}{\|\vec{v}\|}\vec{v}$, which is a unit vector;

$$\left\| \frac{1}{\|\vec{v}\|}\vec{v} \right\| = \frac{1}{\|\vec{v}\|}\|\vec{v}\| = 1. \qquad \text{Note: } \vec{v} \neq \vec{0} \ !$$

The normalized vector points in the same direction, but is now unit length.

ex $\vec{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ $\|\vec{v}\| = \sqrt{1^2+2^2+3^2} = \sqrt{14}$

$$\frac{1}{\|\vec{v}\|}\vec{v} = \begin{bmatrix} 1/\sqrt{14} \\ 2/\sqrt{14} \\ 3/\sqrt{14} \end{bmatrix}$$

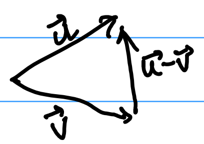
ex Find a basis for $\mathrm{Nul}\begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}$ of unit vectors.

$$= \mathrm{Span}\left\{ \begin{bmatrix} -2 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

norms are $\sqrt{4+1+1} = \sqrt{6}$ and $\sqrt{3}$

so $\left\{ \begin{bmatrix} -2/\sqrt{6} \\ -1/\sqrt{6} \\ 1/\sqrt{6} \\ 0 \end{bmatrix}, \begin{bmatrix} -1/\sqrt{3} \\ -1/\sqrt{3} \\ 0 \\ 1/\sqrt{3} \end{bmatrix} \right\}$

The distance between $\vec{u}, \vec{v} \in \mathbb{R}^n$ is $\mathrm{dist}(\vec{u},\vec{v}) = \|\vec{u}-\vec{v}\|$



ex $\vec{u},\vec{v} \in \mathbb{R}^2$ $\mathrm{dist}(\vec{u},\vec{v}) = \|(u_1,u_2)-(v_1,v_2)\| = \|(u_1-v_1, u_2-v_2)\|$

$$= \sqrt{(u_1-v_1)^2+(u_2-v_2)^2}$$

(the standard Euclidean distance)

$\overbrace{\text{Ortho}}^{\text{right}}\overbrace{\text{gonal}}^{\text{angle}}\text{ity}$

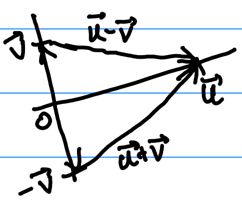
$\vec{u}, \vec{v} \in \mathbb{R}^n$ are **orthogonal** when $\vec{u}\cdot\vec{v} = \vec{0}$

intuition 1: for $\vec{u} = \vec{v}$, $\vec{u}\cdot\vec{v} = \vec{u}\cdot\vec{u} = \|\vec{u}\| \geq \vec{0}$

for $\vec{u} = -\vec{v}$, $\vec{u}\cdot\vec{v} = -\vec{u}\cdot\vec{u} = -\|\vec{u}\| \leq \vec{0}$.

As $\vec{v}$ swings from $\vec{u}$ to $-\vec{u}$, $\vec{u}\cdot\vec{v}$ must be 0 somewhere (intermediate value theorem), and by symmetry it happens "half way" at the right angle.

intuition 2:



$\vec{u}$ and $\vec{v}$ form a right angle when $\|\vec{u}-\vec{v}\| = \|\vec{u}+\vec{v}\|$.

$$\|\vec{u}-\vec{v}\|^2 = (\vec{u}-\vec{v})\cdot(\vec{u}-\vec{v}) = \vec{u}\cdot\vec{u} - 2\vec{u}\cdot\vec{v} + \vec{v}\cdot\vec{v}$$
$$\|\vec{u}+\vec{v}\|^2 = (\vec{u}+\vec{v})\cdot(\vec{u}+\vec{v}) = \vec{u}\cdot\vec{u} + 2\vec{u}\cdot\vec{v} + \vec{v}\cdot\vec{v}$$

These equal exactly when $\vec{u}\cdot\vec{v} = 0$.

intuition 3: Rotating $\begin{pmatrix} a \\ b \end{pmatrix}$ $90°$ CCW gives $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -b \\ a \end{pmatrix}$

$\begin{pmatrix} a \\ b \end{pmatrix}\cdot\begin{pmatrix} -b \\ a \end{pmatrix} = a(-b) + ba = 0$.

Odd consequence of definition: $\vec{0}\cdot\vec{v} = 0$ no matter $\vec{v} \in \mathbb{R}^n$, so $\vec{0}$ is orthogonal to all vectors.

Thm (Pythagorean) $\vec{u}, \vec{v} \in \mathbb{R}^n$ orthogonal $\Leftrightarrow \|\vec{u}\|^2 + \|\vec{v}\|^2 = \|\vec{u}+\vec{v}\|^2$

Pf A previous calculation showed $\|\vec{u}+\vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 + 2\vec{u}\cdot\vec{v}$. $\blacksquare$

(note similarity to law of cosines)
($\vec{u}\cdot\vec{v} = \cos\theta$ for unit vectors)

Orthogonal Complements

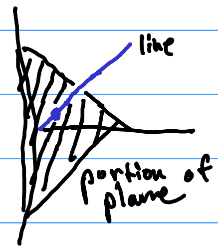
The inner product gives us a way to give a complementary subspace to any subspace of $\mathbb{R}^n$. Let $W \subset \mathbb{R}^n$ be a subspace.

$W^\perp$ ("double-u perp.") is the **orthogonal complement** of W, containing all vectors of $\mathbb{R}^n$ orthogonal to every vector in W.

$$W^\perp = \{ \vec{v} \in \mathbb{R}^n \mid \text{for all } \vec{w} \in W,\ \vec{w}\cdot\vec{v} = 0 \}.$$

ex Let $W = \text{Span}\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}\right\}$. A vector $\vec{v} \in \mathbb{R}^n$ is orthogonal to $\begin{bmatrix}1\\1\\1\end{bmatrix}$ if $0 = \begin{bmatrix}1\\1\\1\end{bmatrix} \cdot \vec{v} = \begin{bmatrix}1 & 1 & 1\end{bmatrix}\vec{v}$. That is, $\vec{v} \in \text{Nul}\begin{bmatrix}1 & 1 & 1\end{bmatrix} = \text{Span}\left\{\begin{bmatrix}-1\\1\\0\end{bmatrix}, \begin{bmatrix}-1\\0\\1\end{bmatrix}\right\}$.

Thus, $W^\perp = \text{Span}\left\{\begin{bmatrix}-1\\1\\0\end{bmatrix}, \begin{bmatrix}-1\\0\\1\end{bmatrix}\right\}$, a plane vs. W a line.



Note: $\begin{bmatrix}c\\c\\c\end{bmatrix} \cdot \vec{v} = c\begin{bmatrix}1\\1\\1\end{bmatrix}\vec{v}$, so being orthogonal to just $\begin{bmatrix}1\\1\\1\end{bmatrix}$ is enough for being orthogonal to everything in W.

ex $\{\vec{0}\}^\perp = \mathbb{R}^n$, since $\vec{v} \cdot \vec{0} = 0$ always.

ex $(\mathbb{R}^n)^\perp = \{\vec{0}\}$ since if $\vec{v} \in (\mathbb{R}^n)^\perp$, $\vec{v} \cdot \vec{x} = 0$ for all $\vec{x} \in \mathbb{R}^n$. One vector $\vec{x}$ may be is $\vec{v}$ itself, in which case $\vec{v} \cdot \vec{v} = 0$. Positive definite implies $\vec{v} = \vec{0}$.

Calculation property If $W = \text{Span}\{\vec{w}_1, \dots, \vec{w}_k\}$, then $\vec{v} \in W^\perp$ if and only if $\vec{v} \cdot \vec{w}_i = 0$ for $1 \le i \le k$. (that is, iff $\vec{v}$ orthogonal to basis of W.)

proof (i) If $\vec{v} \in W^\perp$, then $\vec{v} \cdot \vec{x} = \vec{0}$ for all $\vec{x} \in W$, including when $\vec{x} = \vec{w}_i$.

(ii) If $\vec{v} \cdot \vec{w}_i = 0$ for all i, suppose $\vec{x} \in W$. Let $\vec{c} \in \mathbb{R}^k$ be such that $\vec{x} = c_1\vec{w}_1 + \dots + c_k\vec{w}_k$.
$$\vec{v} \cdot \vec{x} = c_1(\vec{v} \cdot \vec{w}_1) + \dots + c_k(\vec{v} \cdot \vec{w}_k) = c_1 0 + \dots + c_k 0 = 0$$
so $\vec{v} \in W^\perp$. ∎

Prop $W^\perp$ is a subspace of $\mathbb{R}^n$.

pf (i) For $\vec{v}, \vec{w} \in W^\perp$, $\vec{x} \in W$, $(\vec{v} + \vec{w}) \cdot \vec{x} = \vec{v} \cdot \vec{x} + \vec{w} \cdot \vec{x} = 0 + 0 = 0$, so $\vec{v} + \vec{w} \in W^\perp$.

(ii) For $\vec{v} \in W^\perp$, $c \in \mathbb{R}$, $\vec{x} \in W$, $(c\vec{v}) \cdot \vec{x} = c(\vec{v} \cdot \vec{x}) = c\,0 = 0$, so $c\vec{v} \in W^\perp$.

prop $W^\perp \cap W = \{\vec{0}\}$.

pf If $\vec{v} \in W$ and $W^\perp$, $\vec{v} \cdot \vec{x} = 0$ for all $\vec{x} \in W$, including $\vec{x} = \vec{v}$, so $\vec{v} \cdot \vec{v} = 0 \Rightarrow \vec{v} = \vec{0}$ ∎

Relationship to A's subspaces

Let A be $m \times n$

1. $(\text{Row } A)^\perp = \text{Nul } A$.

   Pf If $\vec{x} \in \text{Nul } A$, $A\vec{x} = \vec{0}$. For $A = \begin{bmatrix} - \vec{r}_1 - \\ \vdots \\ - \vec{r}_m - \end{bmatrix}$, $A\vec{x} = \begin{bmatrix} \vec{r}_1 \cdot \vec{x} \\ \vdots \\ \vec{r}_m \cdot \vec{x} \end{bmatrix}$, so $\vec{r}_i \cdot \vec{x} = 0$ for all i. Since Row A is span of rows, and $\vec{x}$ is orthogonal to rows, $\vec{x} \in (\text{Row } A)^\perp$. Conversely, being orthogonal to rows $\Rightarrow$ in Nul A.

2. $(\text{Col } A)^\perp = \text{Nul } A^T$

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

ex $W = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\}$.

$$W^\perp = \left(\text{Col} \begin{bmatrix} 1 & 1 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \right)^\perp = \text{Nul} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 1 \end{bmatrix} = \text{Nul} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix} = \text{Span}\left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Indeed: $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = 0$ and $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = 0$.

ex $$\left(\text{Span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \right)^\perp = \left(\text{Col} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right)^\perp = \text{Nul} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} = \text{Span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

ex If W is k-dim in $\mathbb{R}^n$, $W = \text{Col} \begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_k \end{bmatrix}$.

$W^\perp = \text{Nul} \begin{bmatrix} - \vec{b}_1 - \\ \vdots \\ - \vec{b}_k - \end{bmatrix}$. Since $\begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_k \end{bmatrix}$ has k pivots, and since $\nearrow$ has n columns, it has $n-k$ free columns, so $\dim W^\perp = n-k$.

In fact, $\boxed{\dim W + \dim W^\perp = \dim \mathbb{R}^n}$

prop $(W^\perp)^\perp = W$. pf If $\vec{v} \in W$, $\vec{x} \in W^\perp$, then $\vec{v} \cdot \vec{x} = 0$, so $\vec{v} \in (W^\perp)^\perp$. Thus $W \subset (W^\perp)^\perp$. Since $\dim W^\perp + \dim (W^\perp)^\perp = n$, $\dim W = \dim (W^\perp)^\perp$. ∎

Next time: an orthogonal basis $\vec{b}_1, \ldots, \vec{b}_k$ of W is a basis such that $\vec{b}_i \cdot \vec{b}_j = 0$ when $i \neq j$.

An orthonormal basis in addition has $b_i \cdot b_i = 1$. (normalized)

$$B^{-1} \vec{v} = \begin{bmatrix} \vec{v} \cdot \vec{b}_1 \\ \vdots \\ \vec{v} \cdot \vec{b}_k \end{bmatrix} = \begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_k \end{bmatrix}^T \vec{v}.$$