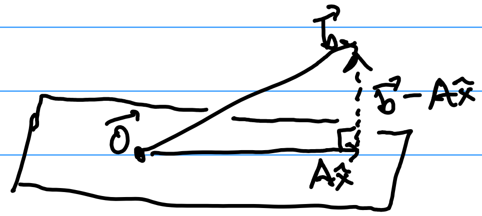
July 20

## Least-Squares

To "solve" an inconsistent system $A\vec{x}=\vec{b}$, we may instead minimize $\|A\vec{x}-\vec{b}\|$ (a square root of a sum of squares, hence "least-squares solution". By the Best Approximation Theorem, this occurs when $A\hat{x}=\text{proj}_{\text{Col }A}\vec{b}$, which is to say when $A\hat{x}-\vec{b}\in(\text{Col }A)^\perp$




Since $(\text{Col }A)^\perp=\text{Nul }A^T$, the condition is $A^T(A\hat{x}-\vec{b})=\vec{0}$, or $A^TA\hat{x}=A^T\vec{b}$.

**Thm** For $m\times n$ $A$, the following are equivalent:

1. $A\vec{x}=\vec{b}$ has a unique least-squares solution for each $\vec{b}$.
2. $\text{rank }A=n$
3. $\text{rank }A^TA=n$ (hence $A^TA$ invertible)

So if any is true, $\hat{x}=(A^TA)^{-1}A^T\vec{b}$ is the unique least-squares solution.

$\|A\hat{x}-\vec{b}\|$ is called the <u>least-squares error</u> or <u>residual</u>.

Alternative methods:

1. If $A$ has orthogonal columns, solve $A\hat{x}=\text{proj}_{\text{Col }A}\vec{b}$ instead.
2. I $A=QR$ is a QR factorization,

$$\text{proj}_{\text{Col }A}\vec{b}=QQ^T\vec{b}\quad(\text{since Col }A=\text{Col }Q)$$

$$QR\hat{x}=QQ^T\vec{b}\Rightarrow Q^TQR\hat{x}=Q^TQQ^T\vec{b}$$

$\Rightarrow R\hat{x}=Q^T\vec{b}$. Use back-substitution or $\hat{x}=R^{-1}Q^T\vec{b}$.

# Inner Product Spaces

We can speak of orthogonality once we have an inner product (like the dot product).

**def** An inner product on a vector space $V$ is a function associating a real number to each pair $\langle \vec{u}, \vec{v} \rangle$, with $\vec{u}, \vec{v} \in V$, which satisfies (for all $\vec{u}, \vec{v}, \vec{w} \in V$ and $c \in \mathbb{R}$),

1. $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$
2. $\langle \vec{u}+\vec{v}, \vec{w} \rangle = \langle \vec{u}, \vec{w} \rangle + \langle \vec{v}, \vec{w} \rangle$
3. $\langle c\vec{u}, \vec{v} \rangle = c\langle \vec{u}, \vec{v} \rangle$
4. $\langle \vec{u}, \vec{u} \rangle \geqslant 0$
5. if $\langle \vec{u}, \vec{u} \rangle = 0$, $\vec{u} = \vec{0}$.

A vector space with a particular inner product is an inner product space

First, let us examine $\langle \vec{u}, \vec{v} \rangle = \vec{u} \cdot \vec{v}$, the standard inner product on $\mathbb{R}^n$. If we forgot vectors had entries,

$$\langle \vec{e}_i, \vec{v} \rangle = v_i$$

so the inner product can recover them. A transformation $T: \mathbb{R}^n \to \mathbb{R}$ is called a dual vector. Its matrix is a $1 \times n$ matrix, say $\vec{w}^T$. Then $T(\vec{x}) = \langle \vec{w}, \vec{x} \rangle$. So every dual vector is an inner product transformation.

In fact, then if $\vec{w} = c_1\vec{e_1} + \cdots + c_n\vec{e_n}$,

$$\langle \vec{w}, \vec{x} \rangle = c_1\langle \vec{e_1}, \vec{x} \rangle + \cdots + c_n\langle \vec{e_n}, \vec{x} \rangle$$
$$= c_1 x_1 + \cdots + c_n x_n$$

But the point is, the dual vector space has dimension $n$, too.

What about matrices? $\langle \vec{y}, A\vec{x} \rangle = \vec{y}^T A \vec{x} = (A^T\vec{y})^T\vec{x}$
$$= \langle A^T\vec{y}, \vec{x} \rangle$$

This is an adjointness relation. Either apply the dual vector for $\vec{y}$ on $A\vec{x}$ or apply the dual vector for $A^T\vec{y}$ on $\vec{x}$.

A need not be square! $\mathbb{R}^m$ inner prod. to $\mathbb{R}^n$ inner prod!

If we want entry $i$ of $A\vec{x}$, $\langle \vec{e_i}, A\vec{x} \rangle = \langle A^T\vec{e_i}, \vec{x} \rangle$
$= \langle \vec{r_i}, \vec{x} \rangle$, where $\vec{r_i}$ is row $i$ of $A$. This is how we compute it already!

Note: we are very close to the idea of "tensor" here, but let's save that for multilinear algebra.

A symmetric matrix $A$ has $A = A^T$. It is self-adjoint:
$$\langle \vec{y}, A\vec{x} \rangle = \vec{y}^T A \vec{x} = \vec{y}^T A^T \vec{x} = \langle A\vec{y}, \vec{x} \rangle.$$

This is important for eigenvectors. Say $A\vec{x} = \lambda\vec{x}$ and $A\vec{y} = \mu\vec{y}$.
$$\langle \vec{y}, A\vec{x} \rangle = \langle \vec{y}, \lambda\vec{x} \rangle = \lambda\langle \vec{y}, \vec{x} \rangle$$
$$\text{〃} \quad \langle A\vec{y}, \vec{x} \rangle = \langle \mu\vec{y}, \vec{x} \rangle = \mu\langle \vec{y}, \vec{x} \rangle$$

so $(\lambda - \mu)\langle \vec{y}, \vec{x} \rangle = 0$. Either $\lambda = \mu$ or $\vec{x}, \vec{y}$ orthogonal.
(this is part of the upcoming spectral theorem).

A common nonstandard inner product on $\mathbb{R}^n$ is
$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^T A \vec{y}$$
where $A$ is $n \times n$ positive definite matrix. That is, $A$ must be a matrix where this is an inner product! Checking $\vec{x}^T A \vec{x} > 0$ for $\vec{x} \neq \vec{0}$ is

sufficient. This means every vector $\vec{x}$ has $A\vec{x}$ within $90°$ of it. So, $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is not positive definite:

$$\vec{x}^T A \vec{x} = [x_1 \; x_2] A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [x_1 \; x_2] \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} = 0$$

If $A$ is diagonal with positive diagonal entries, it is pos. def.
6.7 Ex 1 is $A = \begin{pmatrix} 4 & 0 \\ 0 & 5 \end{pmatrix}$.

Length/norm, distance, projection, orthogonal/orthonormal sets, Gram-Schmidt: these all carry over since they relied only on inner product properties, not dot product properties.

ex $V = \mathbb{P}_2$, $\langle p(x), q(x) \rangle = \left\langle \begin{bmatrix} p(-1) \\ p(0) \\ p(1) \end{bmatrix}, \begin{bmatrix} q(-1) \\ q(0) \\ q(1) \end{bmatrix} \right\rangle$ is pullback of std. inner prod on $\mathbb{R}^3$ to $\mathbb{P}_2$ using evaluation transformation.

Orthogonal basis of $V$?

$1, x, x^2$ basis

$\langle 1, x \rangle = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = 0$ so already orthogonal

$\langle 1, x^2 \rangle = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 2$ $\qquad \langle 1, 1 \rangle = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 3$

$\langle x, x^2 \rangle = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 0$

$x^2 - \frac{2}{3}$ is third vector.

$1, x, x^2 - \frac{2}{3}$ is orthogonal basis.

ex On $C([0,1])$ $\langle f,g\rangle = \int_0^1 f(t)g(t)\,dt$ is an inner product.

For instance,

$$\langle \sin(n\pi x), \sin(m\pi x)\rangle = \int_0^1 \sin(n\pi x)\sin(m\pi x)\,dx$$

$$= \int_0^1 \frac{1}{2}\left(\cos((n-m)\pi x) - \cos((n+m)\pi x)\right)dx$$

if $n=m$, $= \frac{1}{2}\int_0^1 (1-\cos((n+m)\pi x))\,dx$

$$= \frac{1}{2}\left[x - \frac{1}{\pi(n+m)}\sin((n+m)\pi x)\right]_0^1$$

$$= \frac{1}{2}$$

if $n \neq m$ $= \frac{1}{2}\left[\frac{1}{\pi(n-m)}\sin((n-m)\pi x) - \frac{1}{\pi(n+m)}\sin((n+m)\pi x)\right]_0^1$

$= 0.$

So orthogonal.

Can also check $\cos(n\pi x)$ are ortho to these, and to each other.