4. Orthogonality

- Orthogonal sets of vectors
- Orthogonal matrices
- range-nullspace orthogonality

Norms and Inner Products

The norm measures the length of a vector

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{x^T x}$$

It satisfies the *Cauchy-Schwartz* inequality

$$|x^T y| \leq \|x\| \|y\|$$

The angle between two vectors is

$$\theta = \angle(x, y) = \cos^{-1} \frac{x^T y}{\|x\| \|y\|}$$

In particular, x and y are orthogonal if $x^T y = 0$.
Write this as $x \perp y$

Orthonormal set of vectors

Set of vectors $\{u_1, \ldots, u_k\} \in \mathbb{R}^n$ is

- *normalized* if $\|u_i\| = 1$, $i = 1, \ldots, k$
  (u_i are called *unit vectors* or *direction vectors*)
- *orthogonal* if $u_i \perp u_j$ for $i \neq j$
- *orthonormal* if both

slang: we say '$u_1, \ldots, u_k$ are orthonormal vectors' but orthonormality (like independence) is a property of a *set* of vectors, not vectors individually

Orthonormal vectors

Suppose $U = [u_1 \;\; \cdots \;\; u_k]$ has *orthonormal columns*. In matrix notation, this is just

$$U^T U = I$$

Notice this does not mean $UU^T = I$.

Orthonormal vectors are *independent*. Equivalently, $\operatorname{null} U = \{0\}$.

Because

$$\|Ux\|^2 = x^T U^T U x = x^T x = \|x\|^2$$

and so there is no nonzero x such that $Ux = 0$

Geometry

Suppose $U = [u_1 \ \cdots \ u_k]$ has *orthonormal columns*.

- if $w = Uz$ then $\|z\| = \|z\|$

  i.e., U preserves norms. U is called an *isometry*

- if $w_1 = Uz_1$ and $w_2 = Uz_2$ then

  $$w_1^T w_2 = z_1^T z_2$$

  So *angles* are preserved.

- Both properties follow from

  $$w_1^T w_2 = z_1^T U^T U z_2 = z_1^T z_2$$

Orthonormal basis

- Suppose $u_1, \ldots, u_n$ is an orthonormal *basis* for $\mathbb{R}^n$

- then $U = [u_1 \cdots u_n]$ is called **orthogonal**; it is square and satisfies $U^T U = I$

  (you'd think such matrices would be called *orthonormal*, not *orthogonal*)

- it follows that $U^{-1} = U^T$, and hence also $UU^T = I$, *i.e.*,

$$\sum_{i=1}^{n} u_i u_i^T = I$$

Expansion in orthonormal basis

suppose U is orthogonal, so $x = UU^T x$, that is

$$x = \sum_{i=1}^{n} (u_i^T x) u_i$$

- $u_i^T x$ is called the *component* of x in the direction u_i
- $a = U^T x$ *resolves* x into the vector of its u_i components
- $x = Ua$ *reconstitutes* x from its u_i components
- $x = Ua = \sum_{i=1}^{n} a_i u_i$ is called the *expansion* of x

perp notation

- for vectors x, y,

$$x \perp y \qquad \text{means} \qquad x^T y = 0$$

- if S is a subspace

$$x \perp S \qquad \text{means} \qquad x^T y = 0 \text{ for all } y \in S$$

- if S and T are both subspaces

$$T \perp S \qquad \text{means} \qquad x^T y = 0 \text{ for all } x \in T, y \in S$$

- if S is a subspace, the *orthogonal complement* of S

$$S^\perp = \{\, x \mid x \perp S \,\}$$

is the set of all vectors perpendicular to S.

Range and Nullspace Orthogonality

Suppose $A \in \mathbb{R}^{m \times n}$. Then

$$\boxed{(\operatorname{range} A)^\perp = \operatorname{null} A^T}$$

In particular, this implies results quoted in the previous section:

- $\operatorname{range} A = \mathbb{R}^m$ if and only if $\operatorname{null} A^T = \{0\}$
- The columns of A span $\mathbb{R}^m$ if and only if the rows of A are linearly independent.
- $\operatorname{rank} A = \operatorname{rank} A^T$, since

$$\begin{aligned}
\operatorname{rank} A^T &= \dim \operatorname{range} A^T \\
&= \dim (\operatorname{null} A)^\perp \\
&= n - \dim \operatorname{null} A \\
&= \dim \operatorname{range} A \\
&= \operatorname{rank} A
\end{aligned}$$

by conservation of dimension