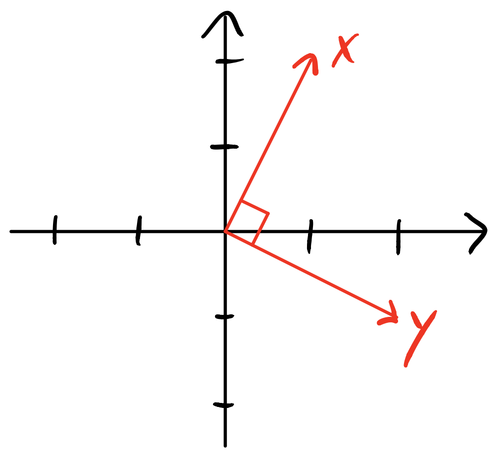
I.5 | Orthogonal Matrices and Subspaces

Example: The vectors $x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $y = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ are orthogonal since $x^T y = 1 \cdot 2 + 2 \cdot (-1) = 0$.



---

① Orthogonal vectors

$\overline{1-2i} = 1+2i$

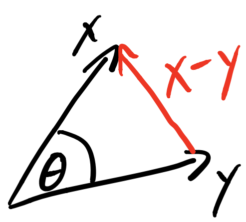
$x, y \in \mathbb{R}^n$ orthogonal if

$$x^T y = x_1 y_1 + \cdots + x_n y_n = 0.$$

$x, y \in \mathbb{C}^n$ orthogonal if

$$x^* y = \bar{x}^T y = \bar{x}_1 y_1 + \cdots + \bar{x}_n y_n = 0.$$

The Law of Cosines



$0 \le \theta \le \pi$

$$\|x-y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\|\|y\|\cos\theta$$

where $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2} = \sqrt{x^T x}$ is the length (norm) of x.

Note that $\|x-y\|^2 = (x-y)^T(x-y)$

$$= (x^T - y^T)(x-y)$$
$$= x^Tx - x^Ty - y^Tx + y^Ty$$
$$= \|x\|^2 - 2x^Ty + \|y\|^2.$$

Therefore, $\boxed{x^Ty = \|x\|\|y\|\cos\theta}$.

Thus, $x^Ty = 0$ iff

$\|x\| = 0$ or $\|y\| = 0$ or $\cos\theta = 0$

(ie $x=0$) (ie $y=0$) (ie $\theta = \pi/2$).

---

② Orthogonal basis

A basis $\{v_1, \dots, v_k\}$ for a subspace in $\mathbb{R}^n$ is othogonal if $v_i^Tv_j = 0$ when $i \neq j$ and is orthonormal if

$$v_i^Tv_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

Note that $v_i^Tv_i = 1$ iff $\|v_i\| = 1$ (ie v_i is a unit vector).

Examples:

(a) The standard basis of $\mathbb{R}^n$

$$e_1, e_2, \ldots, e_n = \begin{bmatrix}1\\0\\0\\\vdots\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\\vdots\\0\\0\end{bmatrix}, \ldots, \begin{bmatrix}0\\0\\0\\\vdots\\0\\1\end{bmatrix}$$

is orthonormal.

(b) A Hadamard matrix is a square ± 1 matrix having othogonal columns.

$$H_2 = \begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}, \quad H_4 = \begin{bmatrix}H_2 & H_2\\H_2 & -H_2\end{bmatrix}, \quad H_8 = \begin{bmatrix}H_4 & H_4\\H_4 & -H_4\end{bmatrix}$$

are Hadamard matrices.

Note that H_2, H_4, and H_8 contain othogonal bases of $\mathbb{R}^2$, $\mathbb{R}^4$, and $\mathbb{R}^8$.

If $\{v_1, \ldots, v_k\}$ is an orthonormal basis for a subspace $S \subseteq \mathbb{R}^n$, then

$$x \in S \Rightarrow x = c_1 v_1 + \cdots + c_k v_k.$$

We can obtain $c_1, \dots, c_k$ by

$$\boxed{c_i = v_i^T x} \quad (i = 1, \dots, k).$$

Pf:

$$\begin{aligned} v_i^T x &= v_i^T (c_1 v_1 + \cdots + c_k v_k) \\ &= c_1 v_i^T v_1 + \cdots + c_i v_i^T v_i + \cdots + c_k v_i^T v_k \\ &= c_1 \cdot 0 + \cdots + c_i \cdot 1 + \cdots + c_k \cdot 0 \\ &= c_i \quad \blacksquare \end{aligned}$$

If $Q = [v_1 \cdots v_k]$, then $Q^T Q = I$,

so $x = Qc \Rightarrow Q^T x = Q^T Q c$

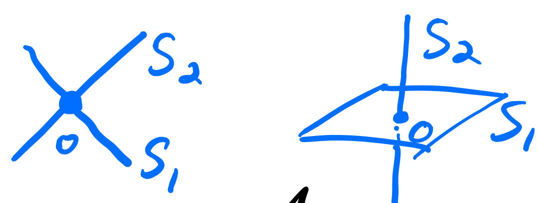
$\Rightarrow Q^T x = I c$

$\Rightarrow c = Q^T x$

(ie $c_i = v_i^T x$, $i = 1, \dots, k$).

---

③ Orthogonal subspaces



The null space of a matrix A is the set of vectors x such that $Ax = 0$.

If $x \in null(A)$ and $y \in col(A^T)$ then $x^T y = 0$.

<u>Pf:</u> Since $y \in col(A^T)$, there is a vector z such that $y = A^T z$.
Then $x^T y = x^T A^T z = (Ax)^T z = 0^T z = 0$. ∎

Therefore, the subspaces $null(A)$ and $col(A^T)$ are <u>orthogonal</u>.
Similarly, the subspaces of $null(A^T)$ and $col(A)$ are <u>orthogonal</u>.

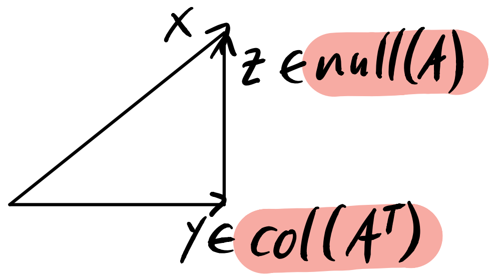
$A \in \mathbb{R}^{m \times n}$ maps vectors in $\mathbb{R}^n$ to vectors in $\mathbb{R}^m$.

$A^T \in \mathbb{R}^{n \times m}$ maps vectors in $\mathbb{R}^m$ to vectors in $\mathbb{R}^n$.

direct sum

$$\mathbb{R}^n = col(A^T) \oplus null(A)$$
$$\mathbb{R}^m = col(A) \oplus null(A^T)$$

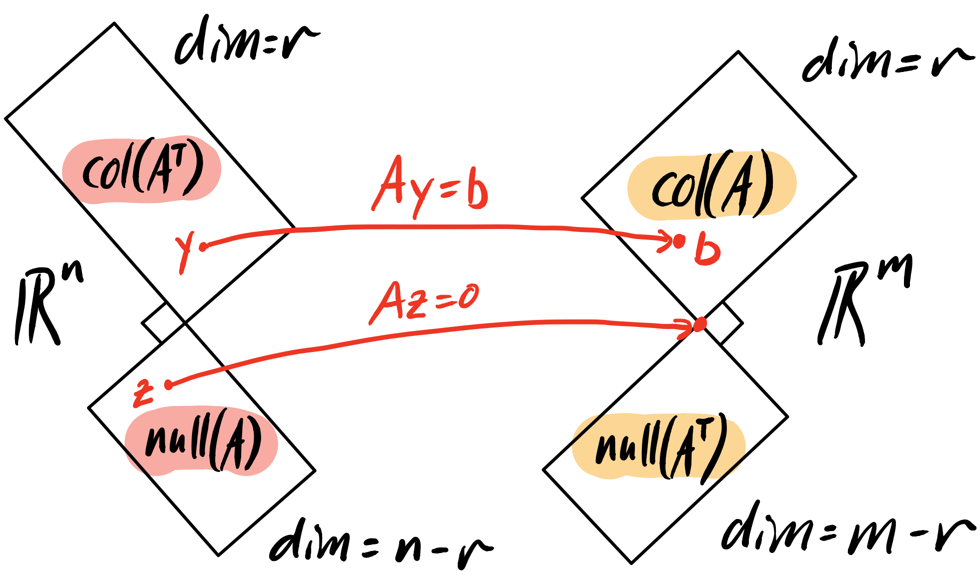
That is, every $x \in \mathbb{R}^n$ can be written as $x = y + z$ where $y \in col(A^T)$ and $z \in null(A)$.



Then

$$Ax = A(y+z) = Ay + Az = Ay + 0 = Ay.$$

Similarly, every vector $x \in \mathbb{R}^m$ can be written as $x = y + z$ where $y \in col(A)$ and $z \in null(A^T)$.



The full SVD of A gives us orthonormal bases of $col(A^T)$, $null(A)$, $col(A)$, and $null(A^T)$.

$$A = U\Sigma V^T \implies \underset{m\times n}{A}\ \underset{n\times n}{V} = \underset{m\times m}{U}\ \underset{m\times n}{\Sigma}$$

$$A\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & & \end{bmatrix} \quad (m \geq n)$$

$\{v_1, \ldots, v_n\}$ is an orthonormal basis of $\mathbb{R}^n$

$\{u_1, \ldots, u_m\}$ is an orthonormal basis of $\mathbb{R}^m$

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$$

$$\sigma_{r+1} = \cdots = \sigma_k = 0 \quad (k = \min(m, n))$$

$$Av_i = \sigma_i u_i \quad (i = 1, \ldots, k)$$

$$\left.\begin{array}{l} \{v_1, \ldots, v_r\} \text{ is a basis of } col(A^T) \\ \{v_{r+1}, \ldots, v_n\} \text{ is a basis of } null(A) \end{array}\right] \mathbb{R}^n$$

$$\left.\begin{array}{l} \{u_1, \ldots, u_r\} \text{ is a basis of } col(A) \\ \{u_{r+1}, \ldots, u_m\} \text{ is a basis of } null(A^T) \end{array}\right] \mathbb{R}^m$$

④ Tall thin Q with orthonormal columns

$$Q = \begin{bmatrix} q_1 \cdots q_k \end{bmatrix} \in \mathbb{R}^{n \times k} \qquad (k \leq n)$$

$$q_i^T q_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \Rightarrow \boxed{Q^T Q = I}$$

But $QQ^T \neq I$ if $k < n$.

Examples:

$$\sqrt{2^2 + 2^2 + (-1)^2} = 3$$

(a) $Q_1 = \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$

$$Q_1^T Q_1 = [1], \quad Q_1 Q_1^T = \frac{1}{9}\begin{bmatrix} 4 & 4 & -2 \\ 4 & 4 & -2 \\ -2 & -2 & 1 \end{bmatrix}$$

(b) $Q_2 = \frac{1}{3}\begin{bmatrix} 2 & 2 \\ 2 & -1 \\ -1 & 2 \end{bmatrix}$

$$Q_2^T Q_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad Q_2 Q_2^T = \frac{1}{9}\begin{bmatrix} 8 & 2 & 2 \\ 2 & 5 & -4 \\ 2 & -4 & 5 \end{bmatrix}$$

(c) $Q_3 = \frac{1}{3}\begin{bmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{bmatrix}$

$Q_3^T Q_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $Q_3 Q_3^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

If Q has orthonormal columns, then $P = QQ^T$ is a <u>projection</u> <u>matrix</u>:

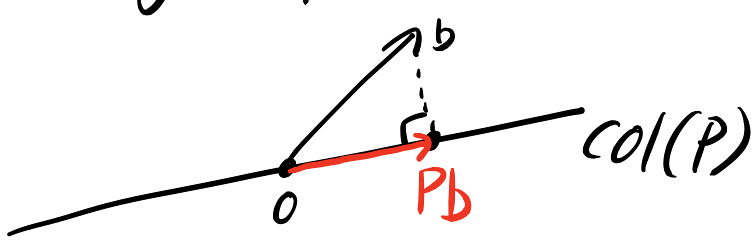
(1) $P^T = P$ (<u>ie</u> P is symmetric)

(2) $P^2 = P$

<u>Pf</u>: (1) $P^T = (QQ^T)^T = (Q^T)^T Q^T = QQ^T = P$

(2) $P^2 = Q(Q^T Q)Q^T = QIQ^T = QQ^T = P$ ∎

Multiplication by P projects a vector orthogonally onto $\mathrm{col}(P) = \mathrm{col}(Q)$.



<u>ie</u> $b - Pb$ is orthog. to $\mathrm{col}(P)$

Examples:

$$b = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$$

(a) $Q_1 = \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$ $\quad P_1 = Q_1 Q_1^T = \frac{1}{9}\begin{bmatrix} 4 & 4 & -2 \\ 4 & 4 & -2 \\ -2 & -2 & 1 \end{bmatrix}$

$P_1 b = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$ $\quad b - P_1 b = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} \perp \operatorname{col}(Q_1)$

(b) $Q_2 = \frac{1}{3}\begin{bmatrix} 2 & 2 \\ 2 & -1 \\ -1 & 2 \end{bmatrix}$ $\quad P_2 = Q_2 Q_2^T = \frac{1}{9}\begin{bmatrix} 8 & 2 & 2 \\ 2 & 5 & -4 \\ 2 & -4 & 5 \end{bmatrix}$

$P_2 b = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}$ $\quad b - P_2 b = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} \perp \operatorname{col}(Q_2)$

---

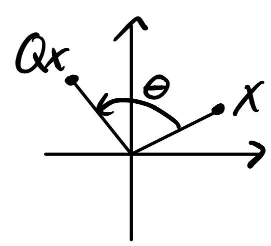
⑤ Orthogonal matrices

Q $n \times n$ and $Q^T Q = Q Q^T = I$

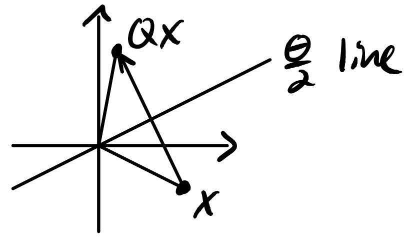
Then $Q^{-1} = Q^T$.

Examples:

(a) $Q_{rotate} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$



(b) $Q_{reflect} = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$



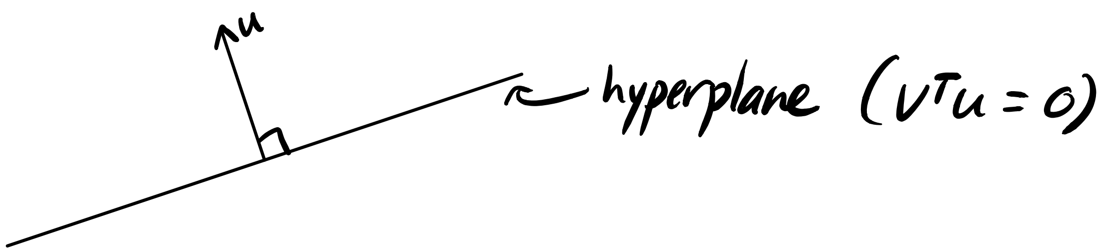
---

* Q_1 and Q_2 orthogonal $\Rightarrow$ Q_1Q_2 orthogonal
* $\|Qx\| = \|x\|$ (length does not change)
* $(Qx)^T(Qy) = x^Ty$ (angles do not change)

Orthogonal matrices are numerically stable.

Householder Reflections

Let $u \in \mathbb{R}^n$ be a unit vector ($\|u\|=1$).
This vector defines a hyperplane:



Then $Q = I - 2uu^T$ is the orthogonal reflector across this hyperplane:

(1) $Qu = -u$

(2) $v^T u = 0 \Rightarrow Qv = v$

(3) $Q^T = Q$

(4) $Q^2 = I$

Pf: $Q^2 = (I - 2uu^T)(I - 2uu^T)$

$= I - 2uu^T - 2uu^T + 4u(u^Tu)u^T$

$= I - 4uu^T + 4u \cdot 1 \cdot u^T$

$= I$ ∎

Example:

$u = \frac{1}{3}\begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} \qquad \|u\| = 1$

$Q = I - 2uu^T = I - \frac{2}{9}\begin{bmatrix} 1 & -2 & -2 \\ -2 & 4 & 4 \\ -2 & 4 & 4 \end{bmatrix} = \frac{1}{9}\begin{bmatrix} 7 & 4 & 4 \\ 4 & 1 & -8 \\ 4 & -8 & 1 \end{bmatrix}$
