Orthogonality and Least Squares

COMP408 - Linear Algebra
Dennis Wong

Magnitude and Direction

The ***magnitude*** of a vector is the distance from the endpoint of the vector to the origin, that is, it's ***length***.

The magnitude of a vector $\vec{a}$, denoted by $|\vec{a}|$, can be computed by the Pythagorean theorem.

Example: $\vec{a} = [4, 3]$ and so $|\vec{a}| = \sqrt{(4^2 + 3^2)} = 5$.

A ***unit vector***, denoted by ^ on top, is a vector of magnitude 1. Unit vectors can be used to express the direction of a vector independent of its magnitude.

Example: The unit vector that corresponds to the direction of $\vec{a} = [4, 3]$ is $\hat{a} = [4, 3] / |\vec{a}| = [4/5, 3/5]$.

Dot Product

A ***dot product*** (or ***inner product***) is the numerical product of the lengths of two vectors, multiplied by the cosine of the angle between them, that is $\vec{a} \cdot \vec{b} = |\vec{a}|\ |\vec{b}|\ cos\ \theta$, where θ represents the angle between the two vectors.

A simply way to calculate a dot product is by multiplying the components of each vector separately and then adding these products together.

Example: $\vec{a} = [4, 3]$, $\vec{b} = [1, 2]$

$\vec{a} \cdot \vec{b} = (4 \text{ x } 3) + (3 \text{ x } 2) = 11$

Orthogonal subspaces

Let V be an inner product space and let S and T be subsets of V . We say that S and T are ***orthogonal***, written $S \perp T$, if every vector in S is orthogonal to every vector in T:

$$S \perp T \text{ iff } s \perp t \text{ for all } s \in S, t \in T.$$

Example: Let $s \in S$ and let $t \in T$. We can write $s = [x_1, x_2, 0]^T$ and $t = [0, 0, x_3]^T$ so that

$$\langle s, t \rangle = s^T t = [x_1, x_2, 0][0, 0, x_3]^T = 0.$$

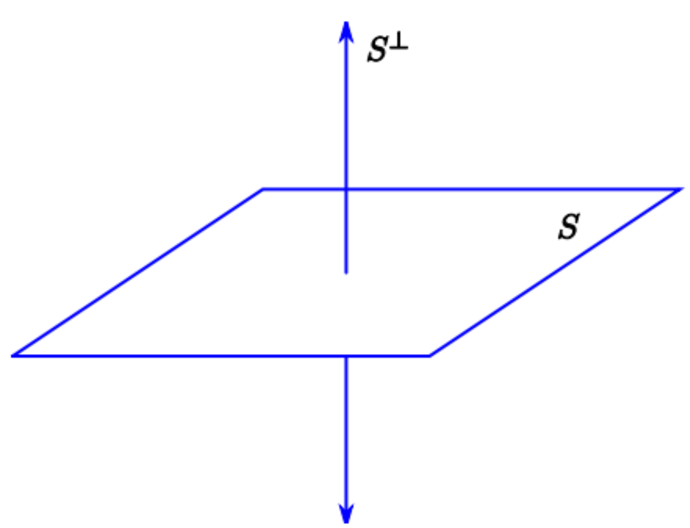
Therefore, $s \perp t$ and we conclude that $S \perp T$.

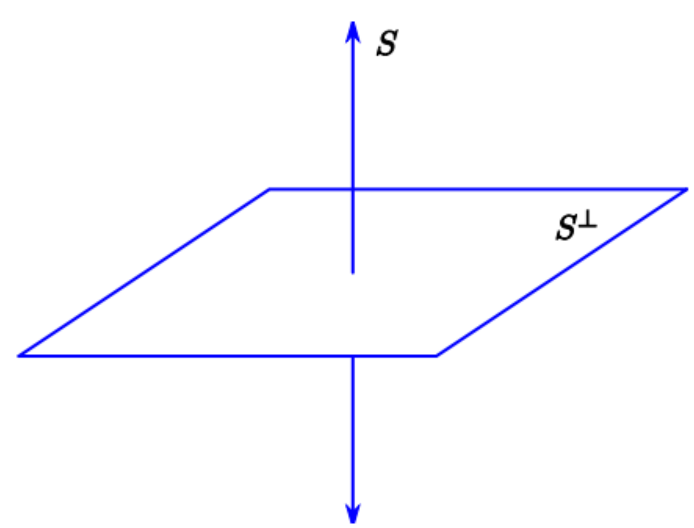
Orthogonal Complement

Let S be a subspace of V. The ***orthogonal complement*** of S (in V), written $S^\perp$, is the set of all vectors in V that are orthogonal to every vector in S:

$S^\perp = \{v \in V \mid v \perp s \text{ for all } s \in S\}$.

Theorem: Let $\{b_1, b_2, \ldots, b_n\}$ be a set vectors in V and let $S = \text{Span}\{b_1, b_2, \ldots, b_n\}$. A vector v in V is in $S^\perp$ if and only if $v \perp b_i$ for each i.

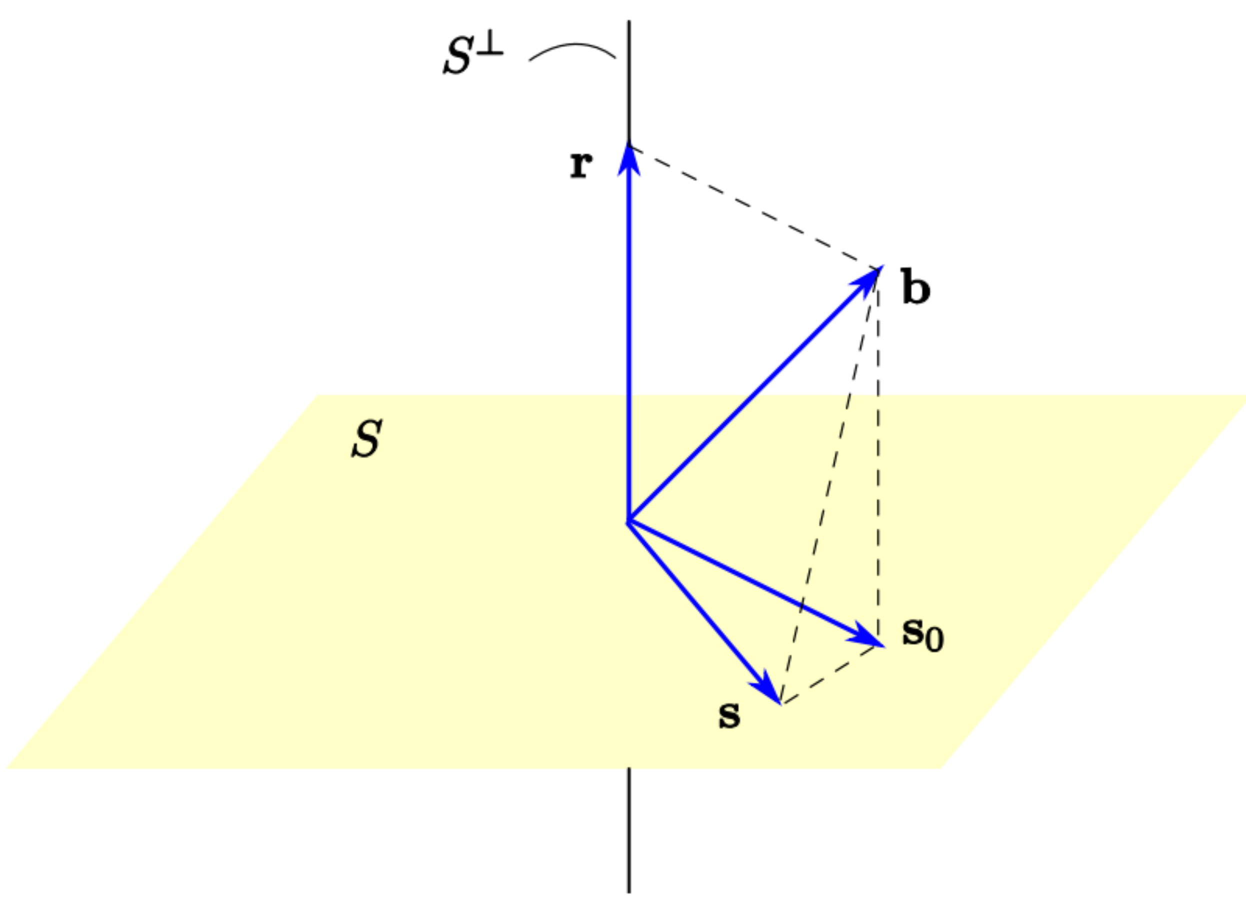




Least squares

Let S be a subspace of V, let b be a vector in V and assume that $b = s_0 + r$ with $s_0 \in$ S and r $\in S^\perp$.

For every $s \in$ S we have dist(b, s_0) $\leq$ dist(b, s).



Least squares

Suppose that the matrix equation $Ax = b$ has no solution. In terms of distance, this means that dist(b, Ax) is never zero, no matter what x is.

Instead of leaving the equation unsolved, it is sometimes useful to find an x_0 that is as close to being a solution as possible.

That is, for which the distance from b to Ax_0 is less than or equal to the distance from b to Ax for every other x. This is called a ***least squares solution***.

Least squares

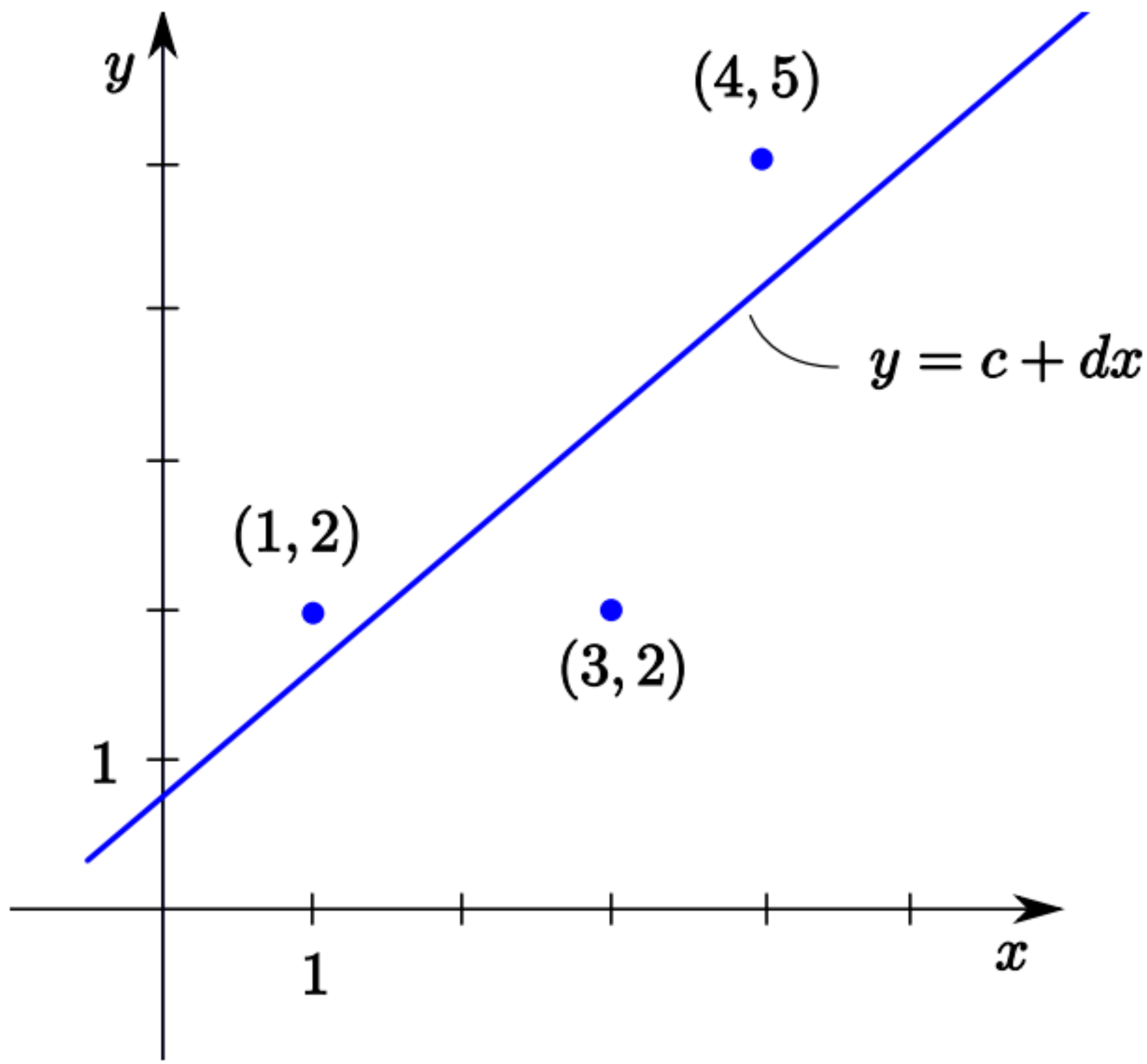
Let A be an $m \times n$ matrix and let b be a vector in R^n. If $x = x_0$ is a solution to

$$A^T A x = A^T b,$$

then, for every $x \in \mathrm{R}^n$, $\mathrm{dist}(b, Ax_0) \leq \mathrm{dist}(b, Ax)$.

Such an x_0 is called a least squares solution to the equation $Ax = b$.

Least squares



Least squares

Example: Use a least squares solution to find a line that best fits the data points (1, 2), (3, 2), and (4, 5).

If we write the desired line as c+dx = y, then ideally the line would go through all three points giving the system

$$\begin{aligned} c + d &= 2 \\ c + 3d &= 2 \\ c + 4d &= 5 \end{aligned}$$

which can be written as the matrix Ax = b with

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} c \\ d \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 2 \\ 2 \\ 5 \end{bmatrix}.$$

Least squares

Applying least square solution by solving the equation $A^{\mathsf{T}}Ax = A^{\mathsf{T}}b,$ we have the following:

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 8 \\ 8 & 26 \end{bmatrix}$$

$$\mathbf{A}^T\mathbf{b} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 9 \\ 28 \end{bmatrix},$$

Least squares

Solving the matrix solution gives us the following:

$$
\begin{array}{cc}
\left[\begin{array}{cc|c} 3 & 8 & 9 \\ 8 & 26 & 28 \end{array}\right] & \begin{array}{c} -8 \\ 3 \end{array}
\end{array}
\sim
\begin{array}{cc}
\left[\begin{array}{cc|c} 3 & 8 & 9 \\ 0 & 14 & 12 \end{array}\right] & \begin{array}{c} \\ \frac{1}{2} \end{array}
\end{array}
$$

$$
\sim
\begin{array}{cc}
\left[\begin{array}{cc|c} 3 & 8 & 9 \\ 0 & 7 & 6 \end{array}\right] & \begin{array}{c} 7 \\ -8 \end{array}
\end{array}
\sim
\begin{array}{cc}
\left[\begin{array}{cc|c} 21 & 0 & 15 \\ 0 & 7 & 6 \end{array}\right] & \begin{array}{c} \frac{1}{21} \\ \frac{1}{7} \end{array}
\end{array}
\sim
\left[\begin{array}{cc|c} 1 & 0 & \frac{5}{7} \\ 0 & 1 & \frac{6}{7} \end{array}\right]
$$

That is, c = 5/7 and d = 6/7 and the best fitting line is y = 5/7 + 6/7x.