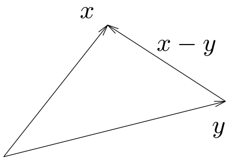
Linear algebra review

- norm, inner product
- subspaces, halfspaces
- independence, basis, dimension
- invertible matrices

(Euclidean) norm

for $x \in \mathbb{R}^n$ we define the (Euclidean) norm as

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



- $\|x\|$ measures length of vector from origin
- $\|x - y\|$ measures distance between x and y
- important properties
  - $\|\alpha x\| = |\alpha| \|x\|$ *homogeneity*
  - $\|x + y\| \le \|x\| + \|y\|$ *triangle inequality*
  - $\|x\| \ge 0$ *nonnegativity*
  - $\|x\| = 0 \iff x = 0$ *definiteness*

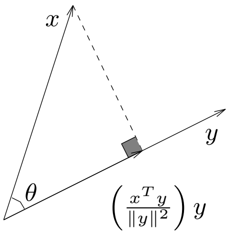
Inner product

$$x^T y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

- called the *dot product* or *scalar product* or *inner product* of x and y, sometimes written $\langle x, y \rangle$
- the Euclidean norm satisfies $||x||^2 = x^T x$
- properties
  - $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$
  - $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
  - $\langle x, y \rangle = \langle y, x \rangle$
  - $\langle x, x \rangle \geq 0$
  - $\langle x, x \rangle = 0 \iff x = 0$

Cauchy-Schwarz inequality and angle between vectors

for any $x, y \in \mathbb{R}^n$



$$\boxed{|x^\mathsf{T} y| \leq ||x|| ||y||}$$

- (unsigned) angle between vectors in $\mathbb{R}^n$ defined as

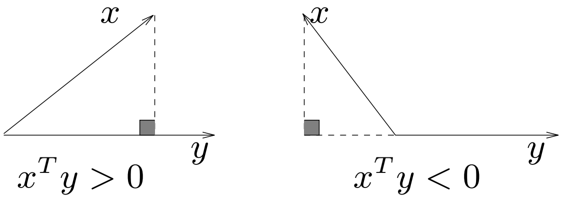
$$\theta = \angle(x, y) = \cos^{-1} \frac{x^\mathsf{T} y}{||x|| ||y||}$$

- thus $x^\mathsf{T} y = ||x|| ||y|| \cos\theta$
- x and y are *aligned*: $\theta = 0$; $x^\mathsf{T} y = ||x|| ||y||$; (if $x \neq 0$) $y = \alpha x$ for some $\alpha \geq 0$
- x and y are *opposed*: $\theta = \pi$; $x^\mathsf{T} y = -||x|| ||y||$ (if $x \neq 0$) $y = -\alpha x$ for some $\alpha \geq 0$
- x and y are *orthogonal*: $\theta = \pi/2$ or $-\pi/2$; $x^\mathsf{T} y = 0$ denoted $x \perp y$

Angles

interpretation of $x^{\mathsf{T}}y > 0$ and $x^{\mathsf{T}}y < 0$

- $x^{\mathsf{T}}y > 0$ means $\angle(x, y)$ is acute
- $x^{\mathsf{T}}y < 0$ means $\angle(x, y)$ is obtuse



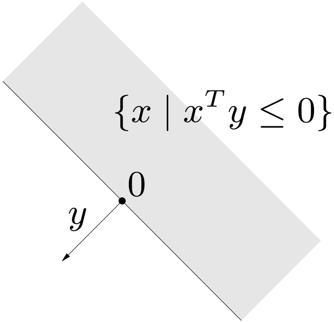
Linear functionals

- a function $f : \mathbb{R}^n \to \mathbb{R}$ is called a *functional*
- for $a \in \mathbb{R}^n$, the function $f(x) = a^\mathsf{T} x$ is a *linear functional*
- every linear functional has this form
- for any $\alpha \in \mathbb{R}$ the set $H_\alpha = \{x \in \mathbb{R}^n \mid a^\mathsf{T} x = \alpha\}$ is called a *hyperplane*
- hyperplanes H_α and H_β are parallel, H_0 passes through the origin

Halfspaces

a *halfspace* with outward normal vector y, and boundary passing through 0

$$H = \{x \mid x^\mathsf{T} y \leq 0\}$$



Subspaces

A set $S \subset \mathbb{R}^n$ is called a subspace if

$$\begin{aligned} x + y \in S \qquad & \text{for all } x, y \in S \\ \alpha x \in S \qquad & \text{for all } \alpha \in \mathbb{R} \text{ and } x \in S \end{aligned}$$

- we say S is *closed under addition* and *closed under scalar multiplication*
- geometrically, S is a flat set which passes through the origin

Examples of subspaces

- $S_1 = \mathbb{R}^n$, *i.e.*, the entire vector space is considered a subspace of itself
- $S_2 = \{0\}$, the origin is the smallest subspace of $\mathbb{R}^n$
- the *span* of a set of vectors $v_1, v_2, \ldots, v_k \in \mathbb{R}^n$ is a subspace

$$\textbf{span}(v_1, v_2, \ldots, v_k) = \{\alpha_1 v_1 + \cdots + \alpha_k v_k \mid \alpha_i \in \mathbb{R}\}$$

  the set of all *linear combinations* of the vectors
- the *sum of two subspaces* is a subspace

$$S + T = \{x + y \mid x \in S, y \in T\}$$

Orthogonal subspaces

- two subspaces $S, T \subset \mathbb{R}^n$ are called *orthogonal* if

$$\boxed{x^\mathsf{T} y = 0 \quad \text{for all } x \in S, y \in T}$$

- for any set $S \subset \mathbb{R}^n$, the *orthogonal complement* is

$$\boxed{S^\perp = \{x \mid x^\mathsf{T} y = 0 \text{ for all } y \in S\}}$$

- $S^\perp$ is always a subspace, even if S is not
- $S^\perp$ is the set of all vectors x, each of which is orthogonal to every vector in S

Independent set of vectors

a set of vectors $\{v_1, v_2, \ldots, v_k\}$ is *independent* if

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0 \quad \Longrightarrow \quad \alpha_1 = \alpha_2 = \cdots = 0$$

some equivalent conditions:

- coefficients of $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k$ are uniquely determined, $i.e.$,

  $$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_k v_k$$

  implies $\alpha_1 = \beta_1, \alpha_2 = \beta_2, \ldots, \alpha_k = \beta_k$
- no vector v_i can be expressed as a linear combination of the other vectors $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k$
- no one vector v_i is in the span of the others

Basis and dimension

set of vectors $\{v_1, v_2, \dots, v_k\}$ is called a *basis* for a subspace S if

$$S = \textbf{span}(v_1, v_2, \dots, v_k)$$

and

$$\{v_1, v_2, \dots, v_k\} \text{ is independent}$$

- equivalently, every $v \in S$ *can be uniquely* expressed as

$$v = \alpha_1 v_1 + \cdots + \alpha_k v_k$$

- for a given subspace S, the number of vectors in any basis is the same, called the *dimension* of S, denoted $d = \textbf{dim}\, S$
- any set of independent vectors in S has no more than d elements
- any set of vectors that span S has at least d elements

Invertibility

A *square* matrix $A \in \mathbb{R}^{n \times n}$ is called *invertible* if there is a matrix $B \in \mathbb{R}^{n \times n}$ such that

$$\boxed{BA = I}$$

- B is called the *inverse* of A, written A^{-1}
- we have $AA^{-1} = A^{-1}A = I$
- A is invertible iff it has linearly independent columns

Interpretations of inverse

suppose $A \in \mathbb{R}^{n \times n}$ has inverse $B = A^{-1}$

- mapping associated with B undoes mapping associated with A (applied either before or after!)
- $x = By$ is a perfect (pre- or post-) *equalizer* for the *channel* $y = Ax$
- $x = By$ is unique solution of $Ax = y$

Change of coordinates

- standard basis vectors in $\mathbb{R}^n$ are $e_1, e_2, \ldots, e_n$, where $e_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$ has a 1 in ith component
- for any x we have
$$x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$$
where x_i are called the *coordinates* of x (in the standard basis)
- if $t_1, t_2, \ldots, t_n$ is another basis for $\mathbb{R}^n$, we have
$$x = \tilde{x}_1 t_1 + \tilde{x}_2 t_2 + \cdots + \tilde{x}_n t_n$$
where $\tilde{x}_i$ are the coordinates of x in the basis $t_1, t_2, \ldots, t_n$
- then $x = T\tilde{x}$ and $\tilde{x} = T^{-1}x$

Similarity transformation

consider linear transformation $y = Ax$, $A \in \mathbb{R}^{n \times n}$

express y and x in terms of $t_1, t_2 \ldots, t_n$, so $x = T\tilde{x}$ and $y = T\tilde{y}$, then

$$\boxed{\tilde{y} = (T^{-1}AT)\tilde{x}}$$

- $A \longrightarrow T^{-1}AT$ is called *similarity transformation*
- similarity transformation by T expresses linear transformation $y = Ax$ in coordinates $t_1, t_2, \ldots, t_n$