Linear functions

Stephen Boyd and Sanjay Lall
Revision by Amirhossein Afsharrad

EE263
Stanford University

Linear equations

Consider system of linear equations:

$$\begin{aligned}
y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\
y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\
&\vdots \\
y_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n
\end{aligned}$$

Can be written in matrix form as $y = Ax$, where

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \quad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Linear functions

$+$ in $\mathbb{R}^n$

$+$ in $\mathbb{R}^m$

A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is *linear* if

- $f(x + y) = f(x) + f(y)$, $\forall x, y \in \mathbb{R}^n$ *(additivity)*
- $f(\alpha x) = \alpha f(x)$, $\forall x \in \mathbb{R}^n$ $\forall \alpha \in \mathbb{R}$ *(homogeneity)*

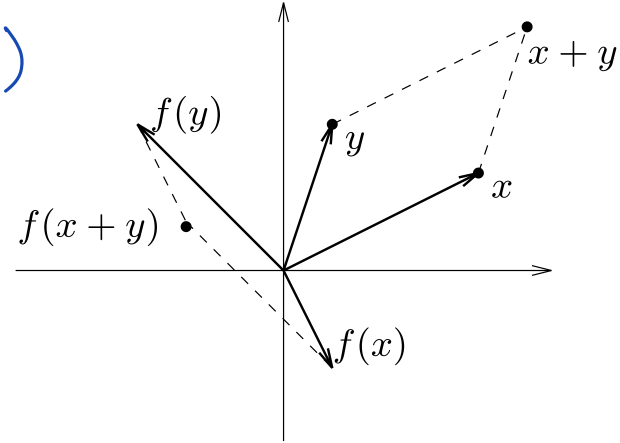
$i.e.$, *superposition* holds

$$f(0) = f(0+0)$$

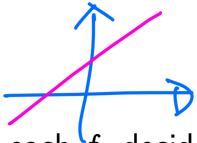
$$= f(0) + f(0)$$

$$f(0) = 2f(0)$$

$$f(0) = 0$$



Exercise: linear or not?



$$x^T y = \underbrace{\begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}}_{1\times n} \underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}}_{n\times 1} = \underbrace{x_1 y_1 + \cdots + x_n y_n}_{\text{inner prod.}}$$

For each f, decide whether it is linear ($a \in \mathbb{R}^n$, $b \in \mathbb{R}$ fixed):

- ✓ ▶ $f(x) = 2x_1 - 3x_2$

$$f(x+y) = 2(x_1+y_1) - 3(x_2+y_2) = \overbrace{(2x_1 - 3x_2)}^{f(x)} + \overbrace{(2y_1 - 3y_2)}^{f(y)}$$

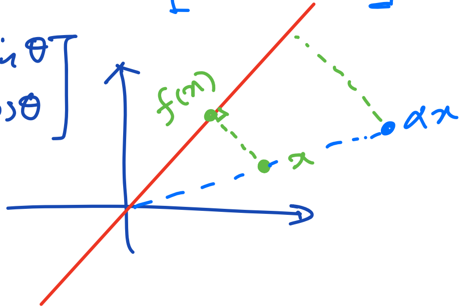
- ✗ ▶ $f(x) = a^{\mathsf{T}} x + b, \quad b \neq 0$

$$f(\alpha x) \neq \alpha f(x)$$

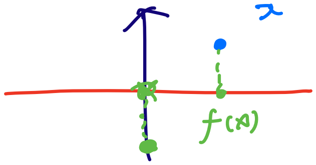
- ✗ ▶ $f(x) = x_1 x_2$
- ✓ ▶ $y_i = x_1 + x_2 + \cdots + x_i$ (running sum)

$$y = Ax, \quad A = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & & 0 \\ 1 & 1 & 1 & 0 & \cdots & \end{bmatrix} \quad n\times n$$

- ✓ ▶ $f : \mathbb{R}^2 \to \mathbb{R}^2$ rotates the plane by an angle θ $\to A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$
- ✓ ▶ $f : \mathbb{R}^2 \to \mathbb{R}^2$ projects x onto the line $y = x$



- ✗ ▶ $f : \mathbb{R}^2 \to \mathbb{R}^2$ projects x onto the line $y = 1$



✓ ▶ $f(x) = Ax$

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

Matrix multiplication function

- Consider function $f : \mathbb{R}^n \to \mathbb{R}^m$ given by $f(x) = Ax$, where $A \in \mathbb{R}^{m \times n}$
- Matrix multiplication function f is linear
- **Converse** is true: **any** linear function $f : \mathbb{R}^n \to \mathbb{R}^m$ can be written as $f(x) = Ax$ for some $A \in \mathbb{R}^{m \times n}$
- Representation via matrix multiplication is unique: for any linear function f there is only one matrix A for which $f(x) = Ax$ for all x
- $y = Ax$ is a concrete representation of a generic linear function

Exercise

$$e_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \rightarrow i\text{th}$$

Show that $f : \mathbb{R}^n \to \mathbb{R}^m$ is linear if and only if $f(x) = Ax$ for some $A \in \mathbb{R}^{m \times n}$.

$$y = \underset{\substack{\downarrow \\ \mathbb{R}^{m\times n}}}{A} x$$

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ x_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \cdots$$

$$= x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$$

$$f(x) = f(x_1 e_1 + \cdots + x_n e_n)$$

$$= f(x_1 e_1) + \cdots + f(x_n e_n)$$

$$= x_1 f(e_1) + \cdots + x_n f(e_n) = \underbrace{\begin{bmatrix} | & & | \\ f(e_1) & \cdots & f(e_n) \\ | & & | \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{x}$$

Interpretations of $y = Ax$

- y is measurement or observation; x is unknown to be determined
- x is 'input' or 'action'; y is 'output' or 'result'
- $y = Ax$ defines a function or transformation that maps $x \in \mathbb{R}^n$ into $y \in \mathbb{R}^m$

Interpretation of a_{ij}

$$y_i = \sum_{j=1}^{n} a_{ij} x_j$$

$y = A x$, $y \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$

a_{ij} is *gain factor* from jth input (x_j) to ith output (y_i)

a_{ij}: i → output, j → input

- ith *row* of A concerns ith *output*
- jth *column* of A concerns jth *input*
- $a_{27} = 0$ means 2nd output (y_2) doesn't depend on 7th input (x_7)
- $|a_{31}| \gg |a_{3j}|$ for $j \neq 1$ means y_3 depends mainly on x_1
- $|a_{52}| \gg |a_{i2}|$ for $i \neq 5$ means x_2 affects mainly y_5
- A is lower triangular, *i.e.*, $a_{ij} = 0$ for $i < j$, means y_i only depends on $x_1, \ldots, x_i$
- A is diagonal, *i.e.*, $a_{ij} = 0$ for $i \neq j$, means ith output depends only on ith input

$$\begin{bmatrix} \times & & 0 & \\ \times & \times & & \\ \times & \times & \times & \\ \times & \times & \times & \times \end{bmatrix}$$

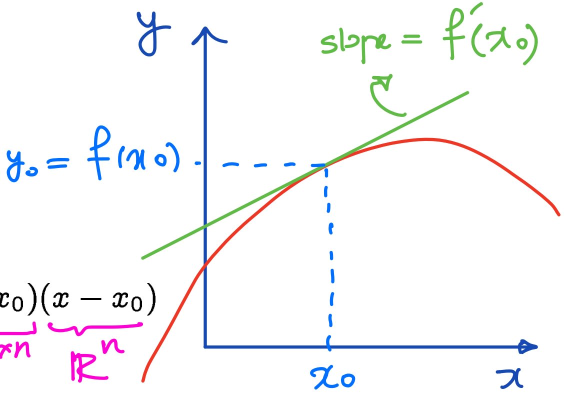
More generally, **sparsity pattern** of A, *i.e.*, list of zero/nonzero entries of A, shows which x_j affect which y_i

Linearization

$$y - y_0 = f'(x_0)(x - x_0)$$

$$y = f'(x_0)(x - x_0) + y_0$$

$$f(x) = f'(x_0)(x - x_0) + f(x_0)$$



- If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $x_0 \in \mathbb{R}^n$, then

$$x \text{ near } x_0 \Longrightarrow \underbrace{f(x)}_{\mathbb{R}^m} \text{ very near } \underbrace{f(x_0)}_{\mathbb{R}^m} + \underbrace{Df(x_0)}_{\mathbb{R}^{m \times n}}\underbrace{(x - x_0)}_{\mathbb{R}^n}$$

  where

$$Df(x_0)_{ij} = \left.\frac{\partial f_i}{\partial x_j}\right|_{x_0}$$

  is derivative (Jacobian) matrix

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 x_2 \\ x_1^2 \end{bmatrix}$$

- With $y = f(x)$, $y_0 = f(x_0)$, define *input deviation* $\delta x := x - x_0$, *output deviation* $\delta y := y - y_0$
- Then we have $\delta y \approx Df(x_0)\delta x$
- When deviations are small, they are (approximately) related by a linear function

$$Df(x) = \begin{bmatrix} x_2 & x_1 \\ 2x_1 & 0 \end{bmatrix}$$