Linear functions

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

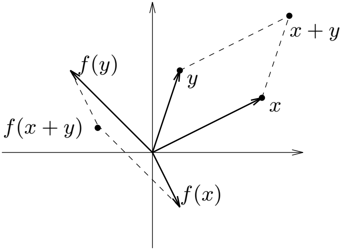
$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \quad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Linear functions

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

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

$i.e.$, *superposition* holds



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

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}

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

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

- 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

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

Linearization

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

  $$x \text{ near } x_0 \Longrightarrow f(x) \text{ very near } f(x_0) + Df(x_0)(x - x_0)$$

  where

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

  is derivative (Jacobian) matrix

- 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