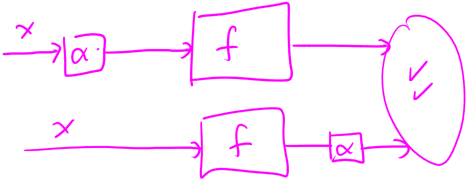
01204211 Discrete Mathematics
Lecture 12b: Linear functions (I)

Jittat Fakcharoenphol

November 12, 2024

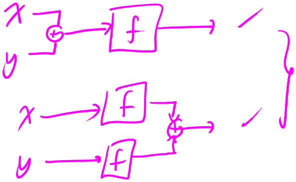
Linear functions



Linear functions

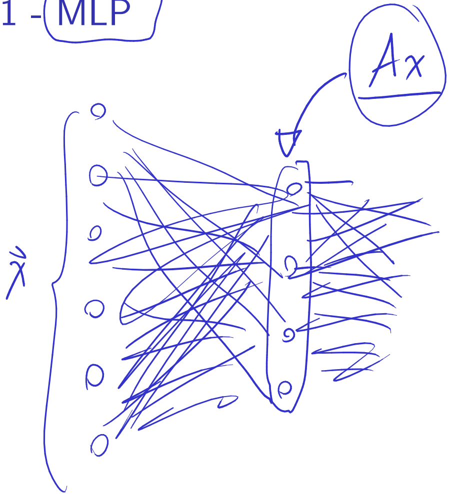
Consider vector spaces $\mathcal{V}$ and $\mathcal{W}$ over $\mathbb{R}$. A function $f : \mathcal{V} \to \mathcal{W}$ is **linear** if

1. for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{V}$, $f(\boldsymbol{x} + \boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$ and
2. for all $\alpha \in \mathbb{R}$ and $\boldsymbol{x} \in \mathcal{V}$, $f(\alpha \boldsymbol{x}) = \alpha f(\boldsymbol{x})$.



$$\boxed{f(x) = Mx}$$

Example 1 - MLP



Example 2 - Page rank (1)

Example 2 - Page rank (2)

Matrix-vector multiplication

Given an $m \times n$ matrix M over $\mathbb{R}$, consider a product

$$M\boldsymbol{x}.$$

Note that for the multiplication to work, $\boldsymbol{x}$ must be in $\mathbb{R}^n$ and the result vector is in $\mathbb{R}^m$. Therefore, we can define a function $f : \mathbb{R}^n \to \mathbb{R}^m$ as

$$f(\boldsymbol{x}) = M\boldsymbol{x}.$$

Note that f is linear because:

$$f(\boldsymbol{x} + \boldsymbol{y}) = M(\boldsymbol{x} + \boldsymbol{y}) = M\boldsymbol{x} + M\boldsymbol{y} = f(\boldsymbol{x}) + f(\boldsymbol{y}),$$

and

$$f(\alpha\boldsymbol{x}) = M(\alpha\boldsymbol{x}) = \alpha M\boldsymbol{x} = \alpha f(\boldsymbol{x}).$$

The converse

Lemma 1

For any linear function $f : \mathbb{R}^n \to \mathbb{R}^m$*, there exists an* $m \times n$ *matrix* M *such that*

$$f(\boldsymbol{x}) = \boxed{M\boldsymbol{x}}.$$

Proof.

Consider any $x \in \mathbb{R}^n$. Let $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]$. Note that

$$\boldsymbol{x} = [x_1, 0, \ldots, 0] + [0, x_2, 0, \ldots, 0] + \cdots + [0, \ldots, 0, x_n].$$

Let $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n \in \mathbb{R}^n$ be standard generators, i.e., $\boldsymbol{e}_i$ be a vector with 1 at the i-th row and 0 at every other positions. (For example $\boldsymbol{e}_1 = [1, 0, \ldots, 0]$ and $\boldsymbol{e}_3 = [0, 0, 1, 0, \ldots, 0]$.)
We thus have

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \cdots + x_n\boldsymbol{e}_n.$$

Since f is linear, this implies that

$$f(\boldsymbol{x}) = x_1 f(\boldsymbol{e}_1) + x_2 f(\boldsymbol{e}_2) + \cdots + x_n f(\boldsymbol{e}_n).$$

$\square$

Proof (cont.)

Define M as follows

$$M = \left[\begin{array}{c|c|c|c} f(\boldsymbol{e}_1) & f(\boldsymbol{e}_2) & \cdots & f(\boldsymbol{e}_n) \end{array} \right].$$

Hence,

$$\begin{aligned} M\boldsymbol{x} &= \left[\begin{array}{c|c|c|c} f(\boldsymbol{e}_1) & f(\boldsymbol{e}_2) & \cdots & f(\boldsymbol{e}_n) \end{array} \right] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= x_1 f(\boldsymbol{e}_1) + x_2 f(\boldsymbol{e}_2) + \cdots + x_n f(\boldsymbol{e}_n) = f(\boldsymbol{x}), \end{aligned}$$

as required. $\square$

Structures of linear functions (overview)