

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 $f : \mathcal{V} \rightarrow \mathcal{W}$ is **linear** if

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

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

$$M\mathbf{x}.$$

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

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

Note that f is linear because:

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

and

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

The converse

Lemma 1

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

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

Proof.

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

$$\mathbf{x} = [x_1, 0, \dots, 0] + [0, x_2, 0, \dots, 0] + \dots + [0, \dots, 0, x_n].$$

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

We thus have

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \dots + x_n\mathbf{e}_n.$$

Since f is linear, this implies that

$$f(\mathbf{x}) = x_1f(\mathbf{e}_1) + x_2f(\mathbf{e}_2) + \dots + x_nf(\mathbf{e}_n).$$



Proof (cont.)

Define M as follows

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

Hence,

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

as required.



Structures of linear functions (overview)