Linear functions and matrices

Linear functions and their representations

Definition 1. (Linear function) *A function $\boldsymbol{f}: \mathbb{R}^N \mapsto \mathbb{R}^M$ is linear if and only if for every $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$ and every $a, b \in \mathbb{R}$,*

$$\boldsymbol{f}(a\,\boldsymbol{x} + b\,\boldsymbol{y}) = a\,\boldsymbol{f}(\boldsymbol{x}) + b\,\boldsymbol{f}(\boldsymbol{y}). \tag{1}$$

Exercise 1. Let $\boldsymbol{f}, \boldsymbol{g}: \mathbb{R}^N \mapsto \mathbb{R}^M$ be linear functions and $a, b \in \mathbb{R}$. Prove that $a\,\boldsymbol{f} + b\,\boldsymbol{g}$ is still a linear function from $\mathbb{R}^N$ to $\mathbb{R}^M$.

Exercise 2. Let $\boldsymbol{f}: \mathbb{R}^N \mapsto \mathbb{R}^M$ be a linear function, and $\boldsymbol{g}: \mathbb{R}^K \mapsto \mathbb{R}^L$ be another linear function. For what N, M, K, L is the composite function $\boldsymbol{f} \circ \boldsymbol{g}$ well-defined? For what N, M, K, L is the composite function $\boldsymbol{g} \circ \boldsymbol{f}$ well-defined? Are they linear? Justify your answer.

Exercise 3. Let $\boldsymbol{f}: \mathbb{R}^N \mapsto \mathbb{R}^M$ be linear. Let $k \in \mathbb{N}$. Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in \mathbb{R}^N$ and $a_1, \ldots, a_k \in \mathbb{R}$. Prove that

$$f(a_1\,\boldsymbol{x}_1 + \cdots + a_k\,\boldsymbol{x}_k) = a_1\,f(\boldsymbol{x}_1) + \cdots + a_k\,f(\boldsymbol{x}_k). \tag{2}$$

In the following we try to obtain a generic formula for linear functions.

The case $M = 1$

Lemma 2. *$f: \mathbb{R}^N \mapsto \mathbb{R}$ is linear $\Longleftrightarrow$ there is $\boldsymbol{a} \in \mathbb{R}^N$ such that $f(\boldsymbol{x}) = \boldsymbol{a} \cdot \boldsymbol{x}$.*

Proof. $\Longleftarrow$ is trivial and left as exercise.

$\Longrightarrow$: Denote by $\boldsymbol{e}_k = (0, \ldots, 0, 1, 0, \ldots, 0)$ where 1 appears only at the k-th position. Then clearly for any $\boldsymbol{x} \in \mathbb{R}^N$,

$$\boldsymbol{x} = (x_1, \ldots, x_N) = x_1\,\boldsymbol{e}_1 + \cdots + x_N\,\boldsymbol{e}_N. \tag{3}$$

Now define

$$\boldsymbol{a} = (f(\boldsymbol{e}_1), \ldots, f(\boldsymbol{e}_N)). \tag{4}$$

We have

$$f(\boldsymbol{x}) = f(x_1\,\boldsymbol{e}_1 + \cdots + x_N\,\boldsymbol{e}_N) = x_1\,f(\boldsymbol{e}_1) + \cdots + x_N\,f(\boldsymbol{e}_N) = \boldsymbol{a} \cdot \boldsymbol{x}. \tag{5}$$

The proof ends. □

General M

Lemma 3. *Let $f: \mathbb{R}^N \mapsto \mathbb{R}^M$ be linear. Denote $f(\boldsymbol{x}) = (f_1(\boldsymbol{x}), \ldots, f_M(\boldsymbol{x}))$. Then each $f_i: \mathbb{R}^N \mapsto \mathbb{R}$ is linear.*

Exercise 4. Prove the above lemma.

Theorem 4. (Representation of linear functions) *Let $f: \mathbb{R}^N \mapsto \mathbb{R}^M$ be linear. Then there are $M\,N$ numbers in $\mathbb{R}$, denoted $a_{11}, a_{12}, \ldots, a_{1N}, a_{21}, \ldots, a_{2N}, \ldots, a_{M1}, \ldots, a_{MN}$, such that*

$$f(\boldsymbol{x}) = (a_{11}\,x_1 + \cdots + a_{1N}\,x_N, \ldots, a_{M1}\,x_1 + \cdots + a_{MN}\,x_N). \tag{6}$$

Proof. This follows immediately from Lemmas 2 and 3. □

Remark 5. (Relation to matrices) If we denote

$$A := \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & & \vdots \\ a_{M1} & \cdots & a_{MN} \end{pmatrix}, \qquad \boldsymbol{x} := \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} \tag{7}$$

then (6) can be re-written into a matrix-vector product:

$$\boldsymbol{f}(\boldsymbol{x}) = A\,\boldsymbol{x}. \tag{8}$$

Remark 6. From the representation formula above, we see that the properties of linear functions are closely related to the properties of matrices. In the following we will review the theory of matrices.

A review of matrices

Definition

Definition 7. *A real $M \times N$ (read: M by N) matrix A is the arrangement of $M\,N$ real numbers into a rectangle of M rows and N columns. That is*

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{MN} \end{pmatrix} \tag{9}$$

Denoted $A \in \mathbb{R}^{M \times N}$.

Remark 8. We usually write $A = (a_{ij})$ with the understanding that a_{ij} is the number at the intersection of the i-th row and the j-th column. For example the “Hilbert matrix”:

$$A = \left(\frac{1}{i+j} \right) \tag{10}$$

means the matrix reads

$$\begin{pmatrix} \frac{1}{1+1} & \frac{1}{1+2} & \cdots \\ \frac{1}{2+1} & \ddots & \\ \vdots & & \ddots \end{pmatrix}. \tag{11}$$

Matrix operations

- Addition/Subtraction: Let $A = (a_{ij})$, $B = (b_{ij})$ be two matrices of the **same size**, then their sum/difference is defined through

$$A \pm B := (a_{ij} \pm b_{ij}). \tag{12}$$

  Exercise 5. Find a matrix $C \in \mathbb{R}^{M \times N}$ such that $C + A = A$ for all $A \in \mathbb{R}^{M \times N}$.

- Scalar multiplication: Let $A = (a_{ij})$, $b \in \mathbb{R}$, then

$$b\,A := (\ b\,a_{ij}\ ). \tag{13}$$

Exercise 6. Find a number $a\in\mathbb{R}$ such that $a\,A=A$ for all $A\in\mathbb{R}^{M\times N}$.

Exercise 7. Check that $\mathbb{R}^{M\times N}$ with the above addition/subtraction and scalar multiplication becomes a real linear vector space. Do you spot any relation between it and $\mathbb{R}^{MN}$?

- Transpose of a matrix: Let $A=(a_{ij})\in\mathbb{R}^{M\times N}$, then its transpose is a matrix $\in\mathbb{R}^{N\times M}$, defined through

$$A^T=\begin{pmatrix} a_{11} & \cdots & a_{M1} \\ \vdots & \ddots & \vdots \\ a_{1N} & \cdots & a_{MN} \end{pmatrix}=(\ a_{ji}\ ). \tag{14}$$

- Matrix multiplication: Let $A=(a_{ij})$ be an $M\times N$ matrix and $B=(b_{ij})$ be an $N\times K$ matrix. Then their matrix product $A\,B$ is defined through

$$A\,B:=\left(\sum_{l=1}^{N} a_{il}\,b_{lj}\right)=a_{i1}\,b_{1j}+a_{i2}\,b_{2j}+\cdots+a_{iN}\,b_{Nj}. \tag{15}$$

Remark 9. Note that the definition of $A\,B$ requires: number of columns of $A=$number of rows of B.

Exercise 8. Let $\boldsymbol{x}=\begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}$ and $\boldsymbol{y}=\begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix}$ be vectors. Prove that $\boldsymbol{x}\cdot\boldsymbol{y}=\boldsymbol{x}^T\boldsymbol{y}$. Note that the left hand side is the inner product in $\mathbb{R}^n$ while the right hand side is matrix multiplication.

Exercise 9. Let $A\in\mathbb{R}^{M\times N}$ and $B\in\mathbb{R}^{N\times K}$. Write $A=\begin{pmatrix} \boldsymbol{a}_1^T \\ \vdots \\ \boldsymbol{a}_M^T \end{pmatrix}$, $B=(\ \boldsymbol{b}_1\ \cdots\ \boldsymbol{b}_N\ )$ where $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_M,\boldsymbol{b}_1,\ldots,\boldsymbol{b}_N\in\mathbb{R}^N$. Prove that $A\,B=(\boldsymbol{a}_i^T\boldsymbol{b}_j)$

Exercise 10. Prove that matrix multiplication is associative: For any A,B,C,

$$(A\,B)\,C=A\,(B\,C) \tag{16}$$

as long as the products are defined. Thus we can simply write the product as $A\,B\,C$.

Exercise 11. Find two matrices A,B such that both $A\,B$ and $B\,A$ are defined but $A\,B\neq B\,A$.

- Matrix-vector multiplication.

  Let $A\in\mathbb{R}^{M\times N}$ and $\boldsymbol{x}\in\mathbb{R}^N$. Then $A\,\boldsymbol{x}$ is defined through treating $\boldsymbol{x}$ as an $N\times 1$ matrix. Thus if $A=\begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{MN} \end{pmatrix}$ and $\boldsymbol{x}=\begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}$, we have

$$A\,\boldsymbol{x}=\begin{pmatrix} a_{11}\,x_1+\cdots+a_{1N}\,x_N \\ \vdots \\ a_{M1}\,x_1+\cdots+a_{MN}\,x_N \end{pmatrix}. \tag{17}$$

Remark 10. (Relation between matrices and vectors) If we denote column vectors:

$$\boldsymbol{a}_i=\begin{pmatrix} a_{1i} \\ \vdots \\ a_{Mi} \end{pmatrix}, \tag{18}$$

then A is a row of such column vectors:

$$A=(\ \boldsymbol{a}_1\ \ \cdots\ \ \boldsymbol{a}_N\ ); \tag{19}$$

We can also denote row vectors

$$\boldsymbol{b}_i^T := \begin{pmatrix} a_{i1} & \cdots & a_{iN} \end{pmatrix} \text{ (that is } \boldsymbol{b}_i := \begin{pmatrix} a_{i1} \\ \vdots \\ a_{iN} \end{pmatrix}\text{)} \tag{20}$$

Then A is a column of such row vectors:

$$A = \begin{pmatrix} \boldsymbol{b}_1^T \\ \vdots \\ \boldsymbol{b}_M^T \end{pmatrix}. \tag{21}$$

With such notation, the transpose of $A = \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_N \end{pmatrix}$ is $\begin{pmatrix} \boldsymbol{a}_1^T \\ \vdots \\ \boldsymbol{a}_N^T \end{pmatrix}$.

Exercise 12. Let A, B be matrices such that $A\,B$ is defined. Prove that $B^T A^T$ is also defined and $B^T A^T = (A\,B)^T$. Explain how this is s generalization of the property of inner product: $\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{y} \cdot \boldsymbol{x}$.

Remark 11. If we write A as a column of row vectors $A = \begin{pmatrix} \boldsymbol{a}_1^T \\ \vdots \\ \boldsymbol{a}_M^T \end{pmatrix}$, then $A\,\boldsymbol{x} = \begin{pmatrix} \boldsymbol{a}_1 \cdot \boldsymbol{x} \\ \vdots \\ \boldsymbol{a}_M \cdot x \end{pmatrix}$, that is each component is an inner product.

Relating linear functions and matrices

Recall that every linear function $\boldsymbol{f}: \mathbb{R}^N \mapsto \mathbb{R}^M$ has the representation:

$$\boldsymbol{f}(\boldsymbol{x}) = A\,\boldsymbol{x} \tag{22}$$

for some matrix $A \in \mathbb{R}^{M \times N}$.

Lemma 12. *Let $\boldsymbol{f}, \boldsymbol{g}: \mathbb{R}^N \mapsto \mathbb{R}^M$ be linear. Let A, B be their matrix representations respectively. Then the the matrix representation for $\boldsymbol{f} \pm \boldsymbol{g}$ is $A \pm B$.*

Proof. We have $(\boldsymbol{f} \pm \boldsymbol{g})(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}) \pm \boldsymbol{g}(\boldsymbol{x}) = A\,\boldsymbol{x} \pm B\,\boldsymbol{x} = (A \pm B)\,\boldsymbol{x}$. □

Lemma 13. *Let $\boldsymbol{f}: \mathbb{R}^N \mapsto \mathbb{R}^M$ be linear with matrix representation $A \in \mathbb{R}^{M \times N}$. Let $a \in \mathbb{R}$. Then the matrix representation for $a\,\boldsymbol{f}$ is $a\,A$.*

Proof. Left as exercise. □

Lemma 14. (Composite functions) *Let $\boldsymbol{f}: \mathbb{R}^n \times \mathbb{R}^m$ and $\boldsymbol{g}: \mathbb{R}^m \mapsto \mathbb{R}^k$ be linear. Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{k \times m}$ be their matrix representations respectively. Then the composite function $\boldsymbol{g} \circ \boldsymbol{f}$ has matrix representation $B\,A$.*

Proof. We have

$$(\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x}) = \boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x})) = \boldsymbol{g}(A\,\boldsymbol{x}) = B\,(A\,\boldsymbol{x}) = (B\,A)\,\boldsymbol{x}. \tag{23}$$

Thus ends the proof. □