3.3 LINEARITY

The concept of linearity is the underlying theme of our subject. In elementary mathematics the term "linear function" refers to straight lines, but in higher mathematics linearity means something much more general. Recall that a function f is simply a rule for associating points in one set $\mathcal{D}$—called the ***domain*** of f—to points in another set $\mathcal{R}$—the ***range*** of f. A *linear* function is a particular type of function that is characterized by the following two properties.

Linear Functions

Suppose that $\mathcal{D}$ and $\mathcal{R}$ are sets that possess an addition operation as well as a scalar multiplication operation—i.e., a multiplication between scalars and set members. A function f that maps points in $\mathcal{D}$ to points in $\mathcal{R}$ is said to be a ***linear function*** whenever f satisfies the conditions that

$$f(x+y) = f(x) + f(y) \tag{3.3.1}$$

and

$$f(\alpha x) = \alpha f(x) \tag{3.3.2}$$

for every x and y in $\mathcal{D}$ and for all scalars α. These two conditions may be combined by saying that f is a linear function whenever

$$f(\alpha x + y) = \alpha f(x) + f(y) \tag{3.3.3}$$

for all scalars α and for all $x, y \in \mathcal{D}$.

One of the simplest linear functions is $f(x) = \alpha x$, whose graph in $\Re^2$ is a straight line through the origin. You should convince yourself that f is indeed a linear function according to the above definition. However, $f(x) = \alpha x + \beta$ does not qualify for the title "linear function"—it is a linear function that has been translated by a constant β. Translations of linear functions are referred to as ***affine functions.*** Virtually all information concerning affine functions can be derived from an understanding of linear functions, and consequently we will focus only on issues of linearity.

In $\Re^3$, the surface described by a function of the form

$$f(x_1, x_2) = \alpha_1 x_1 + \alpha_2 x_2$$

is a plane through the origin, and it is easy to verify that f is a linear function. For $\beta \neq 0$, the graph of $f(x_1, x_2) = \alpha_1 x_1 + \alpha_2 x_2 + \beta$ is a plane *not* passing through the origin, and f is no longer a linear function—it is an affine function.

In $\Re^2$ and $\Re^3$, the graphs of linear functions are lines and planes through the origin, and there seems to be a pattern forming. Although we cannot visualize higher dimensions with our eyes, it seems reasonable to suggest that a general linear function of the form

$$f(x_1, x_2, \ldots, x_n) = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$$

somehow represents a "linear" or "flat" surface passing through the origin $\mathbf{0} = (0, 0, \ldots, 0)$ in $\Re^{n+1}$. One of the goals of the next chapter is to learn how to better interpret and understand this statement.

Linearity is encountered at every turn. For example, the familiar operations of differentiation and integration may be viewed as linear functions. Since

$$\frac{d(f+g)}{dx} = \frac{df}{dx} + \frac{dg}{dx} \quad \text{and} \quad \frac{d(\alpha f)}{dx} = \alpha \frac{df}{dx},$$

the differentiation operator $D_x(f) = df/dx$ is linear. Similarly,

$$\int (f+g)dx = \int f dx + \int g dx \quad \text{and} \quad \int \alpha f dx = \alpha \int f dx$$

means that the integration operator $I(f) = \int f dx$ is linear.

There are several important matrix functions that are linear. For example, the transposition function $f(\mathbf{X}_{m\times n}) = \mathbf{X}^T$ is linear because

$$(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T \quad \text{and} \quad (\alpha \mathbf{A})^T = \alpha \mathbf{A}^T$$

(recall (3.2.1) and (3.2.2)). Another matrix function that is linear is the *trace* function presented below.

Example 3.3.1

The ***trace*** of an $n \times n$ matrix $\mathbf{A} = [a_{ij}]$ is defined to be the sum of the entries lying on the main diagonal of $\mathbf{A}$. That is,

$$trace\,(\mathbf{A}) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii}.$$

Problem: Show that $f(\mathbf{X}_{n\times n}) = trace\,(\mathbf{X})$ is a linear function.

Solution: Let's be efficient by showing that (3.3.3) holds. Let $\mathbf{A} = [a_{ij}]$ and $\mathbf{B} = [b_{ij}]$, and write

$$\begin{aligned} f(\alpha\mathbf{A} + \mathbf{B}) = trace\,(\alpha\mathbf{A} + \mathbf{B}) &= \sum_{i=1}^{n} [\alpha\mathbf{A} + \mathbf{B}]_{ii} = \sum_{i=1}^{n} (\alpha a_{ii} + b_{ii}) \\ &= \sum_{i=1}^{n} \alpha a_{ii} + \sum_{i=1}^{n} b_{ii} = \alpha \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} = \alpha\, trace\,(\mathbf{A}) + trace\,(\mathbf{B}) \\ &= \alpha f(\mathbf{A}) + f(\mathbf{B}). \end{aligned}$$

Example 3.3.2

Consider a linear system

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & u_1, \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & u_2, \\ & & & & \vdots & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & u_m, \end{array}$$

to be a function $\mathbf{u} = f(\mathbf{x})$ that maps $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \Re^n$ to $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix} \in \Re^m$.

Problem: Show that $\mathbf{u} = f(\mathbf{x})$ is linear.

Solution: Let $\mathbf{A} = [a_{ij}]$ be the matrix of coefficients, and write

$$f(\alpha\mathbf{x} + \mathbf{y}) = f\begin{pmatrix} \alpha x_1 + y_1 \\ \alpha x_2 + y_2 \\ \vdots \\ \alpha x_n + y_n \end{pmatrix} = \sum_{j=1}^{n}(\alpha x_j + y_j)\mathbf{A}_{*j} = \sum_{j=1}^{n}(\alpha x_j\mathbf{A}_{*j} + y_j\mathbf{A}_{*j})$$

$$= \sum_{j=1}^{n}\alpha x_j\mathbf{A}_{*j} + \sum_{j=1}^{n} y_j\mathbf{A}_{*j} = \alpha\sum_{j=1}^{n} x_j\mathbf{A}_{*j} + \sum_{j=1}^{n} y_j\mathbf{A}_{*j}$$

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

According to (3.3.3), the function f is linear.

The following terminology will be used from now on.

Linear Combinations

For scalars α_j and matrices $\mathbf{X}_j$, the expression

$$\alpha_1\mathbf{X}_1 + \alpha_2\mathbf{X}_2 + \cdots + \alpha_n\mathbf{X}_n = \sum_{j=1}^{n}\alpha_j\mathbf{X}_j$$

is called a ***linear combination*** of the $\mathbf{X}_j$'s.

Exercises for section 3.3

3.3.1. Each of the following is a function from $\Re^2$ into $\Re^2$. Determine which are linear functions.

(a) $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 1+y \end{pmatrix}.$ (b) $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}.$

(c) $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ xy \end{pmatrix}.$ (d) $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x^2 \\ y^2 \end{pmatrix}.$

(e) $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ \sin y \end{pmatrix}.$ (f) $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ x-y \end{pmatrix}.$

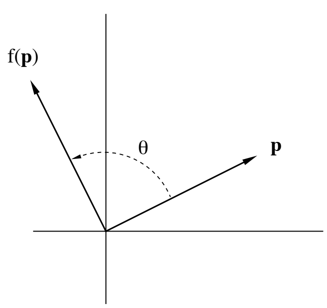
3.3.2. For $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, and for constants ξ_i, verify that

$$f(\mathbf{x}) = \xi_1 x_1 + \xi_2 x_2 + \cdots + \xi_n x_n$$

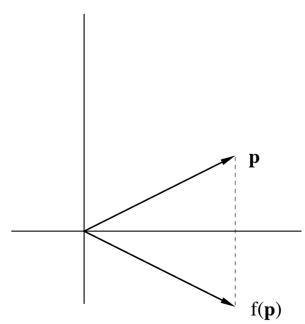
is a linear function.

3.3.3. Give examples of at least two different physical principles or laws that can be characterized as being linear phenomena.

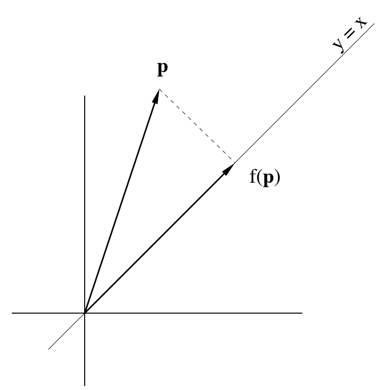
3.3.4. Determine which of the following three transformations in $\Re^2$ are linear.



Rotate counterclockwise through an angle θ.



Reflect about the x-axis.



Project onto the line $y = x$.

Solutions for Chapter 3

Solutions for exercises in section 3. 2

3.2.1. (a) $\mathbf{X}=\begin{pmatrix}0 & 1\\ 2 & 3\end{pmatrix}$ (b) $x=-\frac{1}{2},\ y=-6,$ and $z=0$

3.2.2. (a) Neither (b) Skew symmetric (c) Symmetric (d) Neither

3.2.3. The 3×3 zero matrix trivially satisfies all conditions, and it is the only possible answer for part (a). The only possible answers for (b) are real symmetric matrices. There are many nontrivial possibilities for (c).

3.2.4. $\mathbf{A}=\mathbf{A}^T$ and $\mathbf{B}=\mathbf{B}^T \implies (\mathbf{A}+\mathbf{B})^T=\mathbf{A}^T+\mathbf{B}^T=\mathbf{A}+\mathbf{B}$. Yes—the skew-symmetric matrices are also closed under matrix addition.

3.2.5. (a) $\mathbf{A}=-\mathbf{A}^T \implies a_{ij}=-a_{ji}$. If $i=j$, then $a_{jj}=-a_{jj} \implies a_{jj}=0$.

(b) $\mathbf{A}=-\mathbf{A}^* \implies a_{ij}=-\overline{a_{ji}}$. If $i=j$, then $a_{jj}=-\overline{a_{jj}}$. Write $a_{jj}=x+\mathrm{i}y$ to see that $a_{jj}=-\overline{a_{jj}} \implies x+\mathrm{i}y=-x+\mathrm{i}y \implies x=0 \implies a_{jj}$ is pure imaginary.

(c) $\mathbf{B}^*=(\mathrm{i}\mathbf{A})^*=-\mathrm{i}\mathbf{A}^*=-\mathrm{i}\overline{\mathbf{A}}^T=-\mathrm{i}\mathbf{A}^T=-\mathrm{i}\mathbf{A}=-\mathbf{B}$.

3.2.6. (a) Let $\mathbf{S}=\mathbf{A}+\mathbf{A}^T$ and $\mathbf{K}=\mathbf{A}-\mathbf{A}^T$. Then $\mathbf{S}^T=\mathbf{A}^T+\mathbf{A}^{T^T}=\mathbf{A}^T+\mathbf{A}=\mathbf{S}$. Likewise, $\mathbf{K}^T=\mathbf{A}^T-\mathbf{A}^{T^T}=\mathbf{A}^T-\mathbf{A}=-\mathbf{K}$.

(b) $\mathbf{A}=\frac{\mathbf{S}}{2}+\frac{\mathbf{K}}{2}$ is one such decomposition. To see it is unique, suppose $\mathbf{A}=\mathbf{X}+\mathbf{Y}$, where $\mathbf{X}=\mathbf{X}^T$ and $\mathbf{Y}=-\mathbf{Y}^T$. Thus, $\mathbf{A}^T=\mathbf{X}^T+\mathbf{Y}^T=\mathbf{X}-\mathbf{Y} \implies \mathbf{A}+\mathbf{A}^T=2\mathbf{X}$, so that $\mathbf{X}=\frac{\mathbf{A}+\mathbf{A}^T}{2}=\frac{\mathbf{S}}{2}$. A similar argument shows that $\mathbf{Y}=\frac{\mathbf{A}-\mathbf{A}^T}{2}=\frac{\mathbf{K}}{2}$.

3.2.7. (a) $[(\mathbf{A}+\mathbf{B})^*]_{ij}=[\overline{\mathbf{A}+\mathbf{B}}]_{ji}=[\overline{\mathbf{A}}+\overline{\mathbf{B}}]_{ji}=[\overline{\mathbf{A}}]_{ji}+[\overline{\mathbf{B}}]_{ji}=[\mathbf{A}^*]_{ij}+[\mathbf{B}^*]_{ij}=[\mathbf{A}^*+\mathbf{B}^*]_{ij}$

(b) $[(\alpha\mathbf{A})^*]_{ij}=[\overline{\alpha\mathbf{A}}]_{ji}=[\bar{\alpha}\overline{\mathbf{A}}]_{ji}=\bar{\alpha}[\overline{\mathbf{A}}]_{ji}=\bar{\alpha}[\mathbf{A}^*]_{ij}$

3.2.8. $$k\begin{pmatrix}1 & -1 & 0 & \cdots & 0 & 0\\ -1 & 2 & -1 & \cdots & 0 & 0\\ 0 & -1 & 2 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & 2 & -1\\ 0 & 0 & 0 & \cdots & -1 & 1\end{pmatrix}$$

Solutions for exercises in section 3. 3

3.3.1. Functions (b) and (f) are linear. For example, to check if (b) is linear, let $\mathbf{A}=\begin{pmatrix}a_1\\ a_2\end{pmatrix}$ and $\mathbf{B}=\begin{pmatrix}b_1\\ b_2\end{pmatrix}$, and check if $f(\mathbf{A}+\mathbf{B})=f(\mathbf{A})+f(\mathbf{B})$ and

$f(\alpha\mathbf{A}) = \alpha f(\mathbf{A})$. Do so by writing

$$f(\mathbf{A}+\mathbf{B}) = f\begin{pmatrix} a_1+b_1 \\ a_2+b_2 \end{pmatrix} = \begin{pmatrix} a_2+b_2 \\ a_1+b_1 \end{pmatrix} = \begin{pmatrix} a_2 \\ a_1 \end{pmatrix} + \begin{pmatrix} b_2 \\ b_1 \end{pmatrix} = f(\mathbf{A}) + f(\mathbf{B}),$$

$$f(\alpha\mathbf{A}) = f\begin{pmatrix} \alpha a_1 \\ \alpha a_2 \end{pmatrix} = \begin{pmatrix} \alpha a_2 \\ \alpha a_1 \end{pmatrix} = \alpha\begin{pmatrix} a_2 \\ a_1 \end{pmatrix} = \alpha f(\mathbf{A}).$$

3.3.2. Write $f(\mathbf{x}) = \sum_{i=1}^n \xi_i x_i$. For all points $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, and for all scalars α, it is true that

$$\begin{aligned} f(\alpha\mathbf{x}+\mathbf{y}) &= \sum_{i=1}^n \xi_i(\alpha x_i + y_i) = \sum_{i=1}^n \xi_i \alpha x_i + \sum_{i=1}^n \xi_i y_i \\ &= \alpha \sum_{i=1}^n \xi_i x_i + \sum_{i=1}^n \xi_i y_i = \alpha f(\mathbf{x}) + f(\mathbf{y}). \end{aligned}$$

3.3.3. There are many possibilities. Two of the simplest and most common are Hooke's law for springs that says that $F = kx$ (see Example 3.2.1) and Newton's second law that says that $F = ma$ (i.e., force = mass × acceleration).

3.3.4. They are all linear. To see that rotation is linear, use trigonometry to deduce that if $\mathbf{p} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, then $f(\mathbf{p}) = \mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$, where

$$\begin{aligned} u_1 &= (\cos\theta)x_1 - (\sin\theta)x_2 \\ u_2 &= (\sin\theta)x_1 + (\cos\theta)x_2. \end{aligned}$$

f is linear because this is a special case of Example 3.3.2. To see that reflection is linear, write $\mathbf{p} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $f(\mathbf{p}) = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}$. Verification of linearity is straightforward. For the projection function, use the Pythagorean theorem to conclude that if $\mathbf{p} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, then $f(\mathbf{p}) = \frac{x_1+x_2}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Linearity is now easily verified.