Linear and Affine Maps

Last modified: 2026-11-24 09:19:45 (PST)

Linear and affine maps

Definition 0.1 (Additive function)**.** A function $f : \mathbb{R}^p \to \mathbb{R}^m$ is **additive** [1] if it preserves vector addition: for all $\underset{\sim}{x}, \underset{\sim}{y} \in \mathbb{R}^p$,

$$f(\underset{\sim}{x} + \underset{\sim}{y}) = f(\underset{\sim}{x}) + f(\underset{\sim}{y})$$

Example 0.1 (An additive function, and one that is not)**.**

- $f(x) = 2x$ on $\mathbb{R}$ is additive. For $x, y \in \mathbb{R}$:

$$\begin{aligned} f(x+y) &= 2(x+y) \\ &= 2x + 2y \\ &= f(x) + f(y) \end{aligned}$$

- $g(x) = x + 1$ on $\mathbb{R}$ is not additive, since $g(0+0) = 1$ but $g(0) + g(0) = 2$.

Exercise 0.1 (Testing whether $x_1 x_2$ is additive)**.** An additive function can be evaluated at $\underset{\sim}{x} + \underset{\sim}{y}$ by evaluating it at $\underset{\sim}{x}$ and at $\underset{\sim}{y}$ separately, so additivity is worth checking before splitting a computation that way. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be $f(\underset{\sim}{x}) = x_1 x_2$.
Is f additive?

> **Hint**
>
> Try $\underset{\sim}{x} = (1, 0)$ and $\underset{\sim}{y} = (0, 1)$ in Definition 0.1.

Solution

Solution 0.1. No. For $\underset{\sim}{x} = (1, 0)$ and $\underset{\sim}{y} = (0, 1)$, $\underset{\sim}{x} + \underset{\sim}{y} = (1, 1)$, so

$$f(\underset{\sim}{x} + \underset{\sim}{y}) = 1 \cdot 1 = 1, \qquad f(\underset{\sim}{x}) + f(\underset{\sim}{y}) = 1 \cdot 0 + 0 \cdot 1 = 0.$$

These differ, so f does not satisfy Definition 0.1 for this pair of inputs, and f is not additive.

Definition 0.2 (Homogeneous function)**.** A function $f : \mathbb{R}^p \to \mathbb{R}^m$ is **homogeneous** [1] if it preserves scalar multiplication: for all $\underset{\sim}{x} \in \mathbb{R}^p$ and $c \in \mathbb{R}$,

$$f(c\underset{\sim}{x}) = cf(\underset{\sim}{x})$$

Example 0.2 (Homogeneous functions, and one that is not)**.**

- $f(x) = 2x$ on $\mathbb{R}$ is homogeneous. For $x, c \in \mathbb{R}$:

$$\begin{aligned} f(cx) &= 2cx \\ &= c(2x) \\ &= cf(x) \end{aligned}$$

- $h(x) = x^2$ on $\mathbb{R}$ is not homogeneous, since $h(2 \cdot 1) = 4$ but $2\,h(1) = 2$.
- $k(\underset{\sim}{x}) = (x_1^3 + x_2^3)^{1/3}$ on $\mathbb{R}^2$ is homogeneous but not additive. For $\underset{\sim}{x} \in \mathbb{R}^2$ and $c \in \mathbb{R}$:

$$\begin{aligned} k(c\underset{\sim}{x}) &= (c^3x_1^3 + c^3x_2^3)^{1/3} \\ &= (c^3(x_1^3 + x_2^3))^{1/3} \\ &= c(x_1^3 + x_2^3)^{1/3} \\ &= c\,k(\underset{\sim}{x}) \end{aligned}$$

  but $k((1, 0) + (0, 1)) = k((1, 1)) = 2^{1/3}$, while $k((1, 0)) + k((0, 1)) = 2$.

Exercise 0.2 (Testing whether the absolute value is homogeneous)**.** A homogeneous function turns a rescaled input into the same rescaling of its output, including a rescaling by a negative number. Let $f : \mathbb{R} \to \mathbb{R}$ be the absolute value, $f(x) = |x|$.
Is f homogeneous?

Hint

Try a negative c in Definition 0.2.

Solution

Solution 0.2. No. Take $c = -1$ and $x = 1$. Then

$$f(cx) = |-1| = 1, \qquad cf(x) = -1 \cdot |1| = -1.$$

These differ, so f is not homogeneous. It does satisfy $f(cx) = cf(x)$ for every $c \geq 0$, since then $|cx| = c|x|$.

Definition 0.3 (Linear map, linear transformation, linear function)**.** A function $f : \mathbb{R}^p \to \mathbb{R}^m$ is a **linear map** [1] (also called a **linear transformation** or **linear function**) if it is both additive and homogeneous.

Example 0.3 (A linear map, and two functions that are not)**.**

- $f(\underset{\sim}{x}) = (2x_1,\ x_1 + x_2)$ on $\mathbb{R}^2$ is linear. For $\underset{\sim}{x}, \underset{\sim}{y} \in \mathbb{R}^2$ and $c \in \mathbb{R}$:

$$\begin{aligned} f(\underset{\sim}{x} + \underset{\sim}{y}) &= (2(x_1 + y_1),\ (x_1 + y_1) + (x_2 + y_2)) \\ &= f(\underset{\sim}{x}) + f(\underset{\sim}{y}) \end{aligned}$$

  and

$$\begin{aligned} f(c\underset{\sim}{x}) &= (2cx_1,\ cx_1 + cx_2) \\ &= cf(\underset{\sim}{x}). \end{aligned}$$

- $g(x) = x+1$ on $\mathbb{R}$ is not linear: additivity fails, since $g(0+0) = 1$ but $g(0)+g(0) = 2$.
- $h(x) = x^2$ on $\mathbb{R}$ is not linear: $h(2 \cdot 1) = 4$ but $2\,h(1) = 2$, so $h(cx) \neq c\,h(x)$ for $c = 2$, $x = 1$.

Exercise 0.3 (Testing whether a difference of scaled measurements is a linear map)**.** A score that scales two measurements and takes their difference, such as $3x_1 - x_2$, is a common building block of a model. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be $f(\underset{\sim}{x}) = 3x_1 - x_2$.
Is f a linear map?

Hint

Check the two conditions in Definition 0.3 one at a time, for any $\underset{\sim}{x}, \underset{\sim}{y} \in \mathbb{R}^2$ and $c \in \mathbb{R}$.

Solution

Solution 0.3. Yes. For any $\underset{\sim}{x}, \underset{\sim}{y} \in \mathbb{R}^2$, f is additive:

$$\begin{aligned} f(\underset{\sim}{x} + \underset{\sim}{y}) &= 3(x_1 + y_1) - (x_2 + y_2) && \text{(Definition 9 in Vectors)} \\ &= (3x_1 - x_2) + (3y_1 - y_2) && \text{(rearranging the terms)} \\ &= f(\underset{\sim}{x}) + f(\underset{\sim}{y}). \end{aligned}$$

For any $\underset{\sim}{x} \in \mathbb{R}^2$ and $c \in \mathbb{R}$, f is homogeneous:

$$\begin{aligned} f(c\underset{\sim}{x}) &= 3cx_1 - cx_2 && \text{(Definition 9 in Matrices)} \\ &= c(3x_1 - x_2) && \text{(factoring out } c\text{)} \\ &= cf(\underset{\sim}{x}). \end{aligned}$$

Since f is both, it is a linear map.

Definition 0.4 (Linear operator)**.** A **linear operator** on $\mathbb{R}^p$ is a linear map (Definition 0.3) from $\mathbb{R}^p$ to itself, that is, a linear map $f : \mathbb{R}^p \to \mathbb{R}^p$ whose input and output have the same length. Some authors use "operator" more loosely, for any linear map.

Example 0.4 (A linear operator, and a linear map that is not one)**.** The map $f(\underset{\sim}{x}) = (2x_1,\ x_1 + x_2)$ from Example 0.3 sends $\mathbb{R}^2$ to $\mathbb{R}^2$, so it is a linear operator on $\mathbb{R}^2$. The linear map $g(\underset{\sim}{x}) = x_1 + x_2$ sends $\mathbb{R}^2$ to $\mathbb{R}$, so it is linear but not an operator: its outputs have length 1, not 2.

Exercise 0.4 (Testing whether a projection is a linear operator)**.** Dropping the last coordinate of a point in $\mathbb{R}^3$ projects it onto a plane, as a shadow does. Let $f : \mathbb{R}^3 \to \mathbb{R}^2$ be $f(\underset{\sim}{x}) = (x_1, x_2)$. This f is a linear map.
Is f a linear operator on $\mathbb{R}^3$?

Hint

Write f as $f : \mathbb{R}^p \to \mathbb{R}^m$, and compare p with m as in Definition 0.4.

Solution

Solution 0.4. No. A linear operator on $\mathbb{R}^3$ sends $\mathbb{R}^3$ to $\mathbb{R}^3$. The outputs of f lie in $\mathbb{R}^2$, not $\mathbb{R}^3$, so f is a linear map but not a linear operator.

Theorem 0.1 (A linear map sends zero to zero)**.** *If $f : \mathbb{R}^p \to \mathbb{R}^m$ is a linear map (Definition 0.3), then $f(\underset{\sim}{0}) = \underset{\sim}{0}$.*

Proof

Proof. On the left, $\underset{\sim}{0}$ is the zero vector of length p; on the right, it is the zero vector of length m.

$$\begin{aligned} f(\underset{\sim}{0}) &= f(0 \cdot \underset{\sim}{0}) && \text{(every entry of } 0 \cdot \underset{\sim}{0} \text{ is } 0 \cdot 0 = 0\text{; Definition 9 in Matrices)} \\ &= 0 \cdot f(\underset{\sim}{0}) && \text{(homogeneity, with } c = 0 \text{ and } \underset{\sim}{x} = \underset{\sim}{0}\text{)} \\ &= \underset{\sim}{0} && \text{(every entry of } 0 \cdot f(\underset{\sim}{0}) \text{ is 0; Definition 9 in Matrices)} \end{aligned}$$

□

Example 0.5 (Ruling out linearity at zero)**.** In Example 0.3, $g(x) = x + 1$ has $g(0) = 1 \neq 0$, so by Theorem 0.1 it is not linear, without checking additivity or homogeneity. Sending $\underset{\sim}{0}$ to $\underset{\sim}{0}$ is necessary but not sufficient: $h(x) = x^2$ in the same example has $h(0) = 0$ and is still not linear.

Exercise 0.5 (Ruling out a linear map from its value at zero)**.** Theorem 0.1 gives a quick test: a function that does not send $\underset{\sim}{0}$ to $\underset{\sim}{0}$ cannot be a linear map. Let $g : \mathbb{R}^2 \to \mathbb{R}^2$ be

$$g(\underset{\sim}{x}) = (x_1 + x_2,\ x_2 - 1).$$

Is g a linear map?

Hint

Compute $g(\underset{\sim}{0})$.

Solution

Solution 0.5. No.

$$g(\underset{\sim}{0}) = (0 + 0,\ 0 - 1) = (0, -1) \neq \underset{\sim}{0}.$$

Every linear map sends $\underset{\sim}{0}$ to $\underset{\sim}{0}$ by Theorem 0.1, so g is not one.

Theorem 0.2 (Multiplying by a matrix is a linear map)**.** *For any $m \times p$ matrix $\mathbf{A}$, the function $f : \mathbb{R}^p \to \mathbb{R}^m$ defined by $f(\underset{\sim}{x}) = \mathbf{A}\underset{\sim}{x}$ is a linear map (Definition 0.3).*

Proof

Proof. Both conditions are checked one entry at a time. For each $i \in \{1, \ldots, m\}$, additivity:

$$
\begin{aligned}
(\mathbf{A}(\underset{\sim}{x}+\underset{\sim}{y}))_i &= \sum_{j=1}^{p} a_{ij}(\underset{\sim}{x}+\underset{\sim}{y})_j && \text{(Definition 8 in Matrix Multiplication)}\\
&= \sum_{j=1}^{p} a_{ij}(x_j+y_j) && \text{(Definition 9 in Vectors)}\\
&= \sum_{j=1}^{p} (a_{ij}x_j + a_{ij}y_j) && \text{(multiplication distributes over addition)}\\
&= \sum_{j=1}^{p} a_{ij}x_j + \sum_{j=1}^{p} a_{ij}y_j && \text{(a sum of sums splits)}\\
&= (\mathbf{A}\underset{\sim}{x})_i + (\mathbf{A}\underset{\sim}{y})_i && \text{(Definition 8 in Matrix Multiplication)}\\
&= (\mathbf{A}\underset{\sim}{x} + \mathbf{A}\underset{\sim}{y})_i && \text{(Definition 9 in Vectors)}
\end{aligned}
$$

and homogeneity:

$$
\begin{aligned}
(\mathbf{A}(c\underset{\sim}{x}))_i &= \sum_{j=1}^{p} a_{ij}(c\underset{\sim}{x})_j && \text{(Definition 8 in Matrix Multiplication)}\\
&= \sum_{j=1}^{p} a_{ij}\, c\, x_j && \text{(Definition 9 in Matrices)}\\
&= \sum_{j=1}^{p} c\, a_{ij}x_j && \text{(products are symmetric)}\\
&= c\sum_{j=1}^{p} a_{ij}x_j && \text{(a constant factor comes out of a sum)}\\
&= c\,(\mathbf{A}\underset{\sim}{x})_i && \text{(Definition 8 in Matrix Multiplication)}\\
&= (c\,\mathbf{A}\underset{\sim}{x})_i && \text{(Definition 9 in Matrices)}
\end{aligned}
$$

The four algebra rules cited here are on the algebra page: distributivity, a sum of sums, symmetry of products, and a constant factor. □

Exercise 0.6 (Recognizing a linear map as multiplication by a matrix)**.** Theorem 0.2 saves checking additivity and homogeneity by hand: a function that multiplies its input by a fixed matrix is a linear map. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be

$$f(\underset{\sim}{x}) = (x_1 + 2x_2,\ 3x_1 - x_2).$$

Is f a linear map?

Hint

Find a 2×2 matrix $\mathbf{A}$ with $f(\underset{\sim}{x}) = \mathbf{A}\underset{\sim}{x}$, using Definition 8 in Matrix Multiplication.

Solution

Solution 0.6. Yes. Let

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}.$$

By Definition 8 in Matrix Multiplication,

$$\mathbf{A}\underset{\sim}{x} = \begin{bmatrix} 1 \cdot x_1 + 2 \cdot x_2 \\ 3 \cdot x_1 + (-1) \cdot x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ 3x_1 - x_2 \end{bmatrix} = f(\underset{\sim}{x}).$$

So f multiplies its input by $\mathbf{A}$, and f is a linear map by Theorem 0.2.

Theorem 0.3 (A linear map preserves linear combinations)**.** *If $f : \mathbb{R}^p \to \mathbb{R}^m$ is a linear map (Definition 0.3), then for any vectors $\underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_k \in \mathbb{R}^p$ and any coefficients $c_1, \ldots, c_k \in \mathbb{R}$:*

$$f\left(\sum_{j=1}^{k} c_j \underset{\sim}{v}_j\right) = \sum_{j=1}^{k} c_j f(\underset{\sim}{v}_j)$$

Proof

Proof. By induction on k. For $k = 1$, the claim $f(c_1\underset{\sim}{v}_1) = c_1 f(\underset{\sim}{v}_1)$ is homogeneity. Suppose the claim holds for k vectors. Then for $k+1$:

$$\begin{aligned}
f\left(\sum_{j=1}^{k+1} c_j \underset{\sim}{v}_j\right) &= f\left(\sum_{j=1}^{k} c_j \underset{\sim}{v}_j + c_{k+1}\underset{\sim}{v}_{k+1}\right) && \text{(split off the last term)} \\
&= f\left(\sum_{j=1}^{k} c_j \underset{\sim}{v}_j\right) + f(c_{k+1}\underset{\sim}{v}_{k+1}) && \text{(additivity)} \\
&= \sum_{j=1}^{k} c_j f(\underset{\sim}{v}_j) + f(c_{k+1}\underset{\sim}{v}_{k+1}) && \text{(induction hypothesis)} \\
&= \sum_{j=1}^{k} c_j f(\underset{\sim}{v}_j) + c_{k+1} f(\underset{\sim}{v}_{k+1}) && \text{(homogeneity)} \\
&= \sum_{j=1}^{k+1} c_j f(\underset{\sim}{v}_j) && \text{(join the last term back on)}
\end{aligned}$$

□

Exercise 0.7 (Evaluating a linear map on a linear combination)**.** A linear map is determined on a linear combination of vectors by its values on those vectors alone. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear map with

$$f(\underset{\sim}{v}_1) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \qquad f(\underset{\sim}{v}_2) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

Find $f(3\underset{\sim}{v}_1 - 2\underset{\sim}{v}_2)$.

Hint

Apply Theorem 0.3 with $c_1 = 3$ and $c_2 = -2$.

Solution

Solution 0.7.

$$\begin{aligned} f(3\underset{\sim}{v}_1 - 2\underset{\sim}{v}_2) &= 3f(\underset{\sim}{v}_1) - 2f(\underset{\sim}{v}_2) && \text{(Theorem 0.3)} \\ &= 3\begin{bmatrix} 1 \\ 2 \end{bmatrix} - 2\begin{bmatrix} 0 \\ -1 \end{bmatrix} && \text{(the given values)} \\ &= \begin{bmatrix} 3 \\ 6 \end{bmatrix} - \begin{bmatrix} 0 \\ -2 \end{bmatrix} && \text{(Definition 9 in Matrices)} \\ &= \begin{bmatrix} 3 \\ 8 \end{bmatrix} && \text{(subtracting entry by entry)} \end{aligned}$$

Theorem 0.4 (Every linear map is multiplication by a matrix)**.** *If $f : \mathbb{R}^p \to \mathbb{R}^m$ is a linear map (Definition 0.3), then there is exactly one $m \times p$ matrix $\mathbf{A}$ such that*

$$f(\underset{\sim}{x}) = \mathbf{A}\underset{\sim}{x} \quad \textit{for all } \underset{\sim}{x} \in \mathbb{R}^p$$

Column j of $\mathbf{A}$ is $f(\underset{\sim}{e}_j)$, the value of f at the jth indicator vector (Definition 21 in Vectors); that is, $a_{ij} = \big(f(\underset{\sim}{e}_j)\big)_i$.

Proof

Proof. **Existence.** Let $\mathbf{A}$ be the $m \times p$ matrix with $a_{ij} = \big(f(\underset{\sim}{e}_j)\big)_i$. For any $\underset{\sim}{x} \in \mathbb{R}^p$,

$$
\begin{aligned}
f(\underset{\sim}{x}) &= f\left(\sum_{j=1}^{p} x_j \underset{\sim}{e}_j\right) && \text{(Theorem 4 in Vectors)} \\
&= \sum_{j=1}^{p} x_j f(\underset{\sim}{e}_j) && \text{(Theorem 0.3)}
\end{aligned}
$$

and entry i of that sum is

$$
\begin{aligned}
\left(\sum_{j=1}^{p} x_j f(\underset{\sim}{e}_j)\right)_i &= \sum_{j=1}^{p} x_j \big(f(\underset{\sim}{e}_j)\big)_i && \text{(Definition 11 in Vectors, entry by entry)} \\
&= \sum_{j=1}^{p} x_j a_{ij} && \text{(definition of } \mathbf{A}\text{)} \\
&= \sum_{j=1}^{p} a_{ij} x_j && \text{(products are symmetric)} \\
&= (\mathbf{A}\underset{\sim}{x})_i && \text{(Definition 8 in Matrix Multiplication)}
\end{aligned}
$$

Uniqueness. Suppose $\mathbf{B}$ is any $m \times p$ matrix with $f(\underset{\sim}{x}) = \mathbf{B}\underset{\sim}{x}$ for all $\underset{\sim}{x}$. Taking $\underset{\sim}{x} = \underset{\sim}{e}_j$, for each i:

$$
\begin{aligned}
\big(f(\underset{\sim}{e}_j)\big)_i &= (\mathbf{B}\underset{\sim}{e}_j)_i && \text{(assumption on } \mathbf{B}\text{)} \\
&= \sum_{k=1}^{p} b_{ik} (\underset{\sim}{e}_j)_k && \text{(Definition 8 in Matrix Multiplication)} \\
&= b_{ij} \cdot 1 + \sum_{k \neq j} b_{ik} \cdot 0 && \text{(Definition 21 in Vectors, splitting off the } k = j \text{ term)} \\
&= b_{ij} + 0 && \text{(times 1 changes nothing; times 0 gives 0)} \\
&= b_{ij} && \text{(adding 0 changes nothing)}
\end{aligned}
$$

So

$$
\begin{aligned}
b_{ij} &= \big(f(\underset{\sim}{e}_j)\big)_i \\
&= a_{ij}
\end{aligned}
$$

for every i and j, and $\mathbf{B} = \mathbf{A}$. □

Definition 0.5 (Matrix of a linear map)**.** The **matrix of** a linear map $f : \mathbb{R}^p \to \mathbb{R}^m$ is the unique $m \times p$ matrix $\mathbf{A}$ with $f(\underset{\sim}{x}) = \mathbf{A}\underset{\sim}{x}$ for all $\underset{\sim}{x}$ (Theorem 0.4). Its column j is $f(\underset{\sim}{e}_j)$.

Example 0.6 (Finding the matrix of a linear map)**.** For $f(\underset{\sim}{x}) = (2x_1,\ x_1 + x_2)$ from Example 0.3, the values of f at the indicator vectors are

$$f(\underset{\sim}{e}_1) = f\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}2\cdot 1\\1+0\end{bmatrix} = \begin{bmatrix}2\\1\end{bmatrix}, \qquad f(\underset{\sim}{e}_2) = f\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}2\cdot 0\\0+1\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix}$$

and these are the columns of the matrix of f (Definition 0.5): $\mathbf{A} = \begin{bmatrix}2 & 0\\1 & 1\end{bmatrix}$.

Exercise 0.8 (Evaluating a linear map from its values on the indicator vectors)**.** By Theorem 0.4, the values of a linear map at the indicator vectors determine it everywhere. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear map with

$$f(\underset{\sim}{e}_1) = \begin{bmatrix}1\\3\end{bmatrix}, \qquad f(\underset{\sim}{e}_2) = \begin{bmatrix}-2\\0\end{bmatrix}.$$

Find $f\left(\begin{bmatrix}4\\1\end{bmatrix}\right)$.

Hint

Build the matrix of f from the two given values, then multiply.

Solution

Solution 0.8. By Theorem 0.4, $f(\underset{\sim}{x}) = \mathbf{A}\underset{\sim}{x}$, where column j of $\mathbf{A}$ is $f(\underset{\sim}{e}_j)$:

$$\mathbf{A} = \begin{bmatrix}1 & -2\\3 & 0\end{bmatrix}.$$

So

$$f\left(\begin{bmatrix}4\\1\end{bmatrix}\right) = \begin{bmatrix}1\cdot 4 + (-2)\cdot 1\\3\cdot 4 + 0\cdot 1\end{bmatrix} = \begin{bmatrix}2\\12\end{bmatrix}.$$

Exercise 0.9 (Finding the matrix of a linear map into $\mathbb{R}^3$)**.** The matrix of a linear map turns the map into a table of numbers that software can store and multiply. Let $f : \mathbb{R}^2 \to \mathbb{R}^3$ be the linear map

$$f(\underset{\sim}{x}) = (x_2,\ x_1,\ x_1 - x_2).$$

Find the matrix of f.

Hint

Column j of the matrix is $f(\underset{\sim}{e}_j)$.

Solution

Solution 0.9. The values of f at the indicator vectors are

$$f(\underset{\sim}{e}_1) = f\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}0\\1\\1\end{bmatrix}, \qquad f(\underset{\sim}{e}_2) = f\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}1\\0\\-1\end{bmatrix}.$$

The matrix of f has these values as its columns:

$$\mathbf{A} = \begin{bmatrix}0 & 1\\1 & 0\\1 & -1\end{bmatrix}.$$

It is 3×2, since f sends $\mathbb{R}^2$ to $\mathbb{R}^3$.

Corollary 0.1 (Linear functions from vectors to numbers)**.** *A function $f : \mathbb{R}^p \to \mathbb{R}$ is linear (Definition 0.3) if and only if there is a vector $\underset{\sim}{w} \in \mathbb{R}^p$ such that*

$$f(\underset{\sim}{x}) = \underset{\sim}{w} \cdot \underset{\sim}{x} \quad \text{for all } \underset{\sim}{x} \in \mathbb{R}^p$$

and then $w_j = f(\underset{\sim}{e}_j)$ for each j.

Proof

Proof. **If f is linear.** Theorem 0.4 with $m = 1$ gives a $1 \times p$ matrix $\mathbf{A}$ with $f(\underset{\sim}{x}) = \mathbf{A}\underset{\sim}{x}$ and $a_{1j} = f(\underset{\sim}{e}_j)$. Let $\underset{\sim}{w}$ be the vector with $w_j = a_{1j}$, so that $\mathbf{A} = \underset{\sim}{w}^\top$. Then

$$\begin{aligned} f(\underset{\sim}{x}) &= \mathbf{A}\underset{\sim}{x} && \text{(Theorem 0.4)}\\ &= \underset{\sim}{w} \cdot \underset{\sim}{x} && (\mathbf{A} = \underset{\sim}{w}^\top) \end{aligned}$$

and

$$\begin{aligned} w_j &= a_{1j}\\ &= f(\underset{\sim}{e}_j). \end{aligned}$$

If $f(\underset{\sim}{x}) = \underset{\sim}{w}^\top \underset{\sim}{x}$. Here $\underset{\sim}{w}^\top$ is a $1 \times p$ matrix, so f is linear by Theorem 0.2. □

Example 0.7 (A linear function from $\mathbb{R}^2$ to $\mathbb{R}$)**.** $f(\underset{\sim}{x}) = 3x_1 + 5x_2$ is $\underset{\sim}{w} \cdot \underset{\sim}{x}$ with $\underset{\sim}{w} = (3, 5)^\top$, so it is linear by Corollary 0.1. The entries of $\underset{\sim}{w}$ are the values at the indicator vectors:

$$f(\underset{\sim}{e}_1) = 3 \cdot 1 + 5 \cdot 0 = 3$$

and

$$f(\underset{\sim}{e}_2) = 3 \cdot 0 + 5 \cdot 1 = 5.$$

$g(\underset{\sim}{x}) = 3x_1 + 5x_2 + 1$ is not linear, because $g(\underset{\sim}{0}) = 1 \neq 0$ (Theorem 0.1).

Exercise 0.10 (Finding the weight vector of a linear score)**.** A linear score of p measurements is a dot product of the measurements with a vector of weights, by Corollary 0.1. Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a linear map with

$$f(\underset{\sim}{e}_1) = 2, \qquad f(\underset{\sim}{e}_2) = 0, \qquad f(\underset{\sim}{e}_3) = -1.$$

Find the vector $\underset{\sim}{w}$ with $f(\underset{\sim}{x}) = \underset{\sim}{w} \cdot \underset{\sim}{x}$ for every $\underset{\sim}{x} \in \mathbb{R}^3$.

Hint

Corollary 0.1 gives each entry w_j of $\underset{\sim}{w}$.

Solution

Solution 0.10. By Corollary 0.1, $w_j = f(\underset{\sim}{e}_j)$ for each j, so

$$\underset{\sim}{w} = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}.$$

Then $f(\underset{\sim}{x}) = 2x_1 - x_3$ for every $\underset{\sim}{x} \in \mathbb{R}^3$.

Videos

Two chapters of the *Essence of linear algebra* series by 3Blue1Brown show the results above geometrically:

- Linear transformations and matrices (Chapter 3) shows why a linear map is fixed by where it sends the indicator vectors, which become the columns of its matrix (Definition 0.5).
- Dot products and duality (Chapter 9) shows why every linear map from $\mathbb{R}^p$ to $\mathbb{R}$ is a dot product with a fixed vector (Corollary 0.1).

Definition 0.6 (Translation)**.** For a fixed vector $\underset{\sim}{b} \in \mathbb{R}^m$, the **translation** [2] by $\underset{\sim}{b}$ is the function $t_{\underset{\sim}{b}} : \mathbb{R}^m \to \mathbb{R}^m$ defined by

$$t_{\underset{\sim}{b}}(\underset{\sim}{y}) = \underset{\sim}{y} + \underset{\sim}{b}$$

Example 0.8 (Translating two points)**.** The translation by $\underset{\sim}{b} = (1, -2)^\top$ sends

$$t_{\underset{\sim}{b}}\left(\begin{bmatrix}0\\0\end{bmatrix}\right) = \begin{bmatrix}0+1\\0+(-2)\end{bmatrix} = \begin{bmatrix}1\\-2\end{bmatrix}, \qquad t_{\underset{\sim}{b}}\left(\begin{bmatrix}3\\4\end{bmatrix}\right) = \begin{bmatrix}3+1\\4+(-2)\end{bmatrix} = \begin{bmatrix}4\\2\end{bmatrix}$$

Both points move one unit right and two units down.

Remark (A translation moves every point by the same vector)

Remark 0.1 (A translation moves every point by the same vector)*.* A translation by $\underset{\sim}{b}$ moves every point by the same vector $\underset{\sim}{b}$.

Exercise 0.11 (Finding the translation that moves one point to another)**.** A translation moves every point by the same vector, so where it sends one point fixes the whole translation. A translation $t_{\underset{\sim}{b}}$ of $\mathbb{R}^2$ sends $(2, 3)^\top$ to $(5, 1)^\top$.
Find $\underset{\sim}{b}$.

Hint

Solve $(2, 3)^\top + \underset{\sim}{b} = (5, 1)^\top$ entry by entry.

Solution

Solution 0.11. By Definition 0.6, $t_{\underset{\sim}{b}}\left((2, 3)^\top\right) = (2, 3)^\top + \underset{\sim}{b}$. Matching this vector with $(5, 1)^\top$ entry by entry gives

$$\underset{\sim}{b} = \begin{bmatrix}5-2\\1-3\end{bmatrix} = \begin{bmatrix}3\\-2\end{bmatrix}.$$

Definition 0.7.

Affine map, affine transformation, affine function
A function $f : \mathbb{R}^p \to \mathbb{R}^m$ is an **affine map** (also called an **affine transformation** or **affine function**) if it is a matrix product followed by a translation (Definition 0.6) (a composition, with the matrix product as the inner function): there are an $m \times p$ matrix

$\mathbf{A}$ and a vector $\underset{\sim}{b} \in \mathbb{R}^m$ such that

$$\begin{aligned} f(\underset{\sim}{x}) &= t_{\underset{\sim}{b}}(\mathbf{A}\underset{\sim}{x}) \\ &= \mathbf{A}\underset{\sim}{x} + \underset{\sim}{b} \quad \text{for all } \underset{\sim}{x} \in \mathbb{R}^p \end{aligned}$$

Exercise 0.12 (Testing whether a rescale-and-shift is an affine map)**.** Rescaling a measurement and then shifting it, as when converting units with a different zero point, is the typical affine map. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be

$$f(\underset{\sim}{x}) = (2x_1 + 1,\ x_2 - 3).$$

Is f an affine map?

Hint

Look for a 2×2 matrix $\mathbf{A}$ and a vector $\underset{\sim}{b}$ with $f(\underset{\sim}{x}) = \mathbf{A}\underset{\sim}{x} + \underset{\sim}{b}$.

Solution

Solution 0.12. Yes. Let

$$\mathbf{A} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \underset{\sim}{b} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}.$$

By Definition 8 in Matrix Multiplication, $\mathbf{A}\underset{\sim}{x} = (2x_1, x_2)^\top$, so

$$\mathbf{A}\underset{\sim}{x} + \underset{\sim}{b} = \begin{bmatrix} 2x_1 + 1 \\ x_2 - 3 \end{bmatrix} = f(\underset{\sim}{x})$$

for every $\underset{\sim}{x} \in \mathbb{R}^2$. So f is an affine map.

Definition 0.8 (Offset of an affine map, intercept, bias)**.** In Definition 0.7, the vector $\underset{\sim}{b}$ is called the **offset** of f, also called the **intercept** in statistics, or the **bias** in machine learning.

Remark (Real-valued affine maps)

Remark 0.2 (Real-valued affine maps). For a real-valued function (Definition 6 in Vectors, $m = 1$), $\mathbf{A}$ is a single row, written $\underset{\sim}{w}^\top$, and $\underset{\sim}{b}$ is a single number b, so an affine function is a linear function $\underset{\sim}{w} \cdot \underset{\sim}{x}$ (Corollary 0.1) plus an offset:

$$f(\underset{\sim}{x}) = \underset{\sim}{w} \cdot \underset{\sim}{x} + b, \qquad \underset{\sim}{w} \in \mathbb{R}^p, \quad b \in \mathbb{R} \tag{1}$$

Exercise 0.13 (Reading off the offset of an affine map)**.** In a fitted model, the offset is the part of the output that does not depend on the input. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be the affine map

$$f(\underset{\sim}{x}) = (x_1 - x_2 + 4,\ 3x_2 - 1).$$

Find the offset of f.

Hint

Write $f(\underset{\sim}{x})$ as $\mathbf{A}\underset{\sim}{x} + \underset{\sim}{b}$ as in Definition 0.7.

Solution

Solution 0.13.

$$f(\underset{\sim}{x}) = \begin{bmatrix} 1 & -1 \\ 0 & 3 \end{bmatrix} \underset{\sim}{x} + \begin{bmatrix} 4 \\ -1 \end{bmatrix},$$

so the offset of f is $\underset{\sim}{b} = (4, -1)^\top$. Each entry of $\underset{\sim}{b}$ is the term with no x_j in the matching entry of $f(\underset{\sim}{x})$.

Definition 0.9 (Augmented vector)**.** The **augmented vector** of $\underset{\sim}{x} \in \mathbb{R}^p$ is the vector $\underset{\sim}{x}^{\text{aug}} \in \mathbb{R}^{p+1}$ formed by placing a 1 above the entries of $\underset{\sim}{x}$:

$$\begin{aligned} \underset{\sim}{x}^{\text{aug}} &= \begin{bmatrix} 1 \\ \underset{\sim}{x} \end{bmatrix} \\ &= (1, x_1, \ldots, x_p)^\top \end{aligned}$$

Example 0.9 (An affine function as a dot product with an augmented vector)**.** For $\underset{\sim}{x} = (2, 5)^\top$, the augmented vector is $\underset{\sim}{x}^{\text{aug}} = (1, 2, 5)^\top$. With $\underset{\sim}{w} = (3, -1)^\top$ and $b = 4$, the affine function $\underset{\sim}{w} \cdot \underset{\sim}{x} + b$ (Equation 1) equals the dot product of $(b, w_1, w_2)^\top = (4, 3, -1)^\top$ with $\underset{\sim}{x}^{\text{aug}}$:

$$\begin{aligned} \underset{\sim}{w} \cdot \underset{\sim}{x} + b &= 3 \cdot 2 + (-1) \cdot 5 + 4 \\ &= 5, \\ (4, 3, -1)^\top \cdot \underset{\sim}{x}^{\text{aug}} &= 4 \cdot 1 + 3 \cdot 2 + (-1) \cdot 5 \\ &= 5 \end{aligned}$$

Exercise 0.14 (Absorbing an offset into a dot product with the augmented vector)**.** The augmented vector lets software treat an affine score as a dot product, with the offset as

one more weight. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be

$$f(\underset{\sim}{x}) = 2x_1 - x_2 + 5.$$

Find the vector $\underset{\sim}{v} \in \mathbb{R}^3$ with $f(\underset{\sim}{x}) = \underset{\sim}{v} \cdot \underset{\sim}{x}^{\text{aug}}$ for every $\underset{\sim}{x} \in \mathbb{R}^2$.

Hint

Follow Example 0.9. The first entry of $\underset{\sim}{x}^{\text{aug}}$ is 1.

Solution

Solution 0.14. Let $\underset{\sim}{v} = (5, 2, -1)^\top$. Since $\underset{\sim}{x}^{\text{aug}} = (1, x_1, x_2)^\top$ by Definition 0.9,

$$\underset{\sim}{v} \cdot \underset{\sim}{x}^{\text{aug}} = 5 \cdot 1 + 2x_1 + (-1)x_2 = f(\underset{\sim}{x}).$$

The first entry of $\underset{\sim}{v}$ is the offset 5, and the other two are the numbers multiplying x_1 and x_2.

Theorem 0.5 (An affine map sends zero to its offset)**.** *An affine map $f(\underset{\sim}{x}) = \mathbf{A}\underset{\sim}{x} + \underset{\sim}{b}$ sends the zero vector to its offset $\underset{\sim}{b}$:*

$$f(\underset{\sim}{0}) = \underset{\sim}{b}. \tag{2}$$

Proof

Proof. First, $\mathbf{A}\underset{\sim}{0} = \underset{\sim}{0}$: for each i,

$$\begin{aligned}
(\mathbf{A}\underset{\sim}{0})_i &= \sum_{j=1}^{p} a_{ij} \cdot 0 && \text{(Definition 8 in Matrix Multiplication)} \\
&= \sum_{j=1}^{p} 0 && \text{(any number times 0 is 0)} \\
&= 0 && \text{(a sum of zeros is 0)}
\end{aligned}$$

Then

$$\begin{aligned}
f(\underset{\sim}{0}) &= \mathbf{A}\underset{\sim}{0} + \underset{\sim}{b} && \text{(Definition 0.7)} \\
&= \underset{\sim}{0} + \underset{\sim}{b} && (\mathbf{A}\underset{\sim}{0} = \underset{\sim}{0}) \\
&= \underset{\sim}{b} && \text{(adding } \underset{\sim}{0} \text{ changes no entry; Definition 9 in Vectors)}
\end{aligned}$$

□

Exercise 0.15 (Recovering a column of an affine map's matrix from two values)**.** Theorem 0.5 says that the value of an affine map at $\underset{\sim}{0}$ is its offset, so subtracting that value isolates the matrix part. Let $f(\underset{\sim}{x}) = \mathbf{A}\underset{\sim}{x} + \underset{\sim}{b}$ be an affine map from $\mathbb{R}^2$ to $\mathbb{R}^2$ with

$$f(\underset{\sim}{0}) = \begin{bmatrix}1\\2\end{bmatrix}, \qquad f\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}4\\2\end{bmatrix}.$$

Find the first column of $\mathbf{A}$.

Hint

$\mathbf{A}\begin{bmatrix}1\\0\end{bmatrix}$ is the first column of $\mathbf{A}$.

Solution

Solution 0.15. By Theorem 0.5, $\underset{\sim}{b} = f(\underset{\sim}{0}) = (1,2)^\top$. So

$$\begin{aligned}\mathbf{A}\begin{bmatrix}1\\0\end{bmatrix} &= f\left(\begin{bmatrix}1\\0\end{bmatrix}\right) - \underset{\sim}{b} && \text{(Definition 0.7)}\\ &= \begin{bmatrix}4-1\\2-2\end{bmatrix} && \text{(the given values)}\\ &= \begin{bmatrix}3\\0\end{bmatrix}.\end{aligned}$$

By Definition 8 in Matrix Multiplication, $\mathbf{A}(1,0)^\top$ is the first column of $\mathbf{A}$, which is therefore $(3,0)^\top$.

Theorem 0.6 (An affine map is linear exactly when its offset is zero)**.** *An affine map $f(\underset{\sim}{x}) = \mathbf{A}\underset{\sim}{x} + \underset{\sim}{b}$ is linear if and only if its offset is zero:*

$$f \text{ is linear} \iff \underset{\sim}{b} = \underset{\sim}{0}. \tag{3}$$

Proof

Proof. If $\underset{\sim}{b} = \underset{\sim}{0}$, then

$$\begin{aligned}f(\underset{\sim}{x}) &= \mathbf{A}\underset{\sim}{x} + \underset{\sim}{0}\\ &= \mathbf{A}\underset{\sim}{x},\end{aligned}$$

which is linear by Theorem 0.2. Conversely, if f is linear, then $f(\underset{\sim}{0}) = \underset{\sim}{0}$ by Theorem 0.1, and $f(\underset{\sim}{0}) = \underset{\sim}{b}$ by Theorem 0.5, so $\underset{\sim}{b} = \underset{\sim}{0}$. □

Example 0.10 (Two affine maps with the same matrix)**.** Let $\mathbf{A} = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$, the matrix of the linear map in Example 0.3. $f(\underset{\sim}{x}) = \mathbf{A}\underset{\sim}{x}$ has offset $\underset{\sim}{0}$, so it is linear. $g(\underset{\sim}{x}) = \mathbf{A}\underset{\sim}{x} + \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ has offset $\begin{bmatrix} 1 \\ 0 \end{bmatrix} \neq \underset{\sim}{0}$, so by Theorem 0.6 it is not linear: indeed $g(\underset{\sim}{0}) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, not $\underset{\sim}{0}$.

Exercise 0.16 (Choosing a constant that makes an affine map linear)**.** By Theorem 0.6, an affine map is a linear map exactly when its offset is $\underset{\sim}{0}$. For $c \in \mathbb{R}$, let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be

$$f(\underset{\sim}{x}) = (x_1 - x_2 + c - 2,\ 3x_1 + c^2 - 4).$$

For which values of c is f linear?

Hint

Find the offset of f as a function of c.

Solution

Solution 0.16. Only for $c = 2$. The map is

$$f(\underset{\sim}{x}) = \begin{bmatrix} 1 & -1 \\ 3 & 0 \end{bmatrix} \underset{\sim}{x} + \begin{bmatrix} c - 2 \\ c^2 - 4 \end{bmatrix},$$

so its offset is $(c - 2,\ c^2 - 4)^\top$. By Theorem 0.6, f is linear exactly when both entries are 0. The first entry is 0 only for $c = 2$, and then the second entry is $2^2 - 4 = 0$ too.

Corollary 0.2 (Every linear map is affine)**.** *Every linear map $f : \mathbb{R}^p \to \mathbb{R}^m$ is an affine map with offset $\underset{\sim}{0}$. With $\mathbf{A}$ the matrix of f,*

$$f(\underset{\sim}{x}) = \mathbf{A}\underset{\sim}{x} + \underset{\sim}{0} \quad \text{for all } \underset{\sim}{x} \in \mathbb{R}^p. \tag{4}$$

Proof

Proof. If f is linear, Theorem 0.4 gives a matrix $\mathbf{A}$ with

$$\begin{aligned} f(\underset{\sim}{x}) &= \mathbf{A}\underset{\sim}{x} \\ &= \mathbf{A}\underset{\sim}{x} + \underset{\sim}{0}, \end{aligned}$$

which is an affine map with offset $\underset{\sim}{b} = \underset{\sim}{0}$. □

Exercise 0.17 (Testing whether a linear score is an affine map)**.** Some results hold only for affine maps, so it helps to check whether a given score qualifies. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be $f(\underset{\sim}{x}) = x_1 - 4x_2$, a linear map.
Is f an affine map?

Hint

Use Corollary 0.2.

Solution

Solution 0.17. Yes. By Corollary 0.2, every linear map is an affine map with offset $\underset{\sim}{0}$. Here

$$f(\underset{\sim}{x}) = \begin{bmatrix} 1 & -4 \end{bmatrix} \underset{\sim}{x} + 0,$$

with the 1×2 matrix $\begin{bmatrix} 1 & -4 \end{bmatrix}$ and offset 0.

Theorem 0.7 (An affine map determines its matrix and offset)**.** *If an affine map f satisfies $f(\underset{\sim}{x}) = \mathbf{A}\underset{\sim}{x} + \underset{\sim}{b}$ and $f(\underset{\sim}{x}) = \mathbf{A}'\underset{\sim}{x} + \underset{\sim}{b}'$ for all $\underset{\sim}{x} \in \mathbb{R}^p$, then the two formulas have the same matrix and the same offset:*

$$\mathbf{A}' = \mathbf{A} \quad \text{and} \quad \underset{\sim}{b}' = \underset{\sim}{b}. \tag{5}$$

Proof

Proof. By Theorem 0.5, applied to each formula,

$$\begin{aligned} \underset{\sim}{b} &= f(\underset{\sim}{0}) \\ &= \underset{\sim}{b}'. \end{aligned}$$

Then

$$\begin{aligned} \mathbf{A}\underset{\sim}{x} &= f(\underset{\sim}{x}) - \underset{\sim}{b} \\ &= \mathbf{A}'\underset{\sim}{x} \end{aligned}$$

for all $\underset{\sim}{x}$, so $\mathbf{A}$ and $\mathbf{A}'$ are both matrices (Definition 0.5) of the linear map $\underset{\sim}{x} \mapsto f(\underset{\sim}{x}) - \underset{\sim}{b}$ (linear by Theorem 0.2), and the uniqueness in Theorem 0.4 gives $\mathbf{A}' = \mathbf{A}$. □

Exercise 0.18 (Matching the offsets of two formulas for one affine map)**.** Theorem 0.7 says that an affine map has only one matrix and one offset, so two ways of writing the same map must agree term by term. For $c \in \mathbb{R}$, suppose that

$$\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \underset{\sim}{x} + \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \underset{\sim}{x} + \begin{bmatrix} c \\ 1 \end{bmatrix} \quad \text{for every } \underset{\sim}{x} \in \mathbb{R}^2.$$

Find c.

Hint

Compare the offsets of the two sides.

Solution

Solution 0.18. $c = 3$. Both sides describe the same affine map. By Theorem 0.7, the two sides have the same offset, so $(3, 1)^\top = (c, 1)^\top$ and $c = 3$. Setting $\underset{\sim}{x} = \underset{\sim}{0}$ gives the same answer directly.

Remark ("Linear" in this site, and in machine learning and statistics)

Remark 0.3 ("Linear" in this site, and in machine learning and statistics). This site uses "linear" only in the sense of Definition 0.3, so by Theorem 0.6 a function with a nonzero offset is affine but not linear. For functions of one variable, this is the same split as affine $ax + b$ versus linear ax on the algebra page. (The algebra page writes these as $mx + b$ and mx; here m already names the number of outputs.)
Machine learning and statistics often call any function of the form Equation 1 "linear", as in a *linear model* in statistics or a *linear layer* in a neural network, and call the offset b the bias (machine learning) or the intercept (statistics), as Definition 0.7 notes. A model with an intercept is affine in $\underset{\sim}{x}$ but linear in its coefficients $(b, \underset{\sim}{w})$, and that is the sense of "linear" in "linear model": it stays linear in the coefficients even when the inputs are transformed: $w_1 x + w_2 x^2 + b$ is not affine in x, but it is linear in (b, w_1, w_2). Separately, the affine function $\underset{\sim}{w} \cdot \underset{\sim}{x} + b$ is a linear function of the augmented vector $\underset{\sim}{x}^{\text{aug}}$ (Definition 0.9), with b as the coefficient on its first entry.

Further reading

- Zhou [3], sections 6 ("Linear functions and operators"; see Definition 0.4) and 7 ("Affine functions"), proves that every linear function is a matrix-vector product by the same route as Theorem 0.4, and that every affine function has the form $\mathbf{A}\underset{\sim}{x} + \underset{\sim}{b}$. It defines "affine" differently, by $f(\alpha \underset{\sim}{x} + \beta y) = \alpha f(\underset{\sim}{x}) + \beta f(y)$ whenever $\alpha + \beta = 1$, and shows that this condition holds exactly for the maps $\mathbf{A}\underset{\sim}{x} + \underset{\sim}{b}$, so the two definitions agree.
- Boyd and Vandenberghe [4], chapter 2 ("Linear functions"), covers the vector-to-number case: linear functions as inner products, affine functions, the first-order Taylor approximation as an affine function, and the regression model. Section 8.1

("Linear and affine functions") covers the vector-to-vector case. A free PDF is on the book's website.

Definition 0.10 (Hyperplane)**.** A **hyperplane** [5] in $\mathbb{R}^p$ is the set of points $\underset{\sim}{x} \in \mathbb{R}^p$ satisfying

$$\underset{\sim}{w} \cdot \underset{\sim}{x} + b = 0$$

for some vector $\underset{\sim}{w} \in \mathbb{R}^p \setminus \{\underset{\sim}{0}\}$ and number $b \in \mathbb{R}$. The number b is the offset of the hyperplane.

Definition 0.11 (Normal vector of a hyperplane)**.** In Definition 0.10, the vector $\underset{\sim}{w}$ is a **normal vector** [6] of the hyperplane.

Example 0.11 (A line in $\mathbb{R}^2$, and why the normal vector must be nonzero)**.**

- With $\underset{\sim}{w} = (1, 1)$ and $b = -1$, the hyperplane in $\mathbb{R}^2$ is the line $x_1 + x_2 - 1 = 0$. The point $(1, 0)$ is on it, since $1+0-1 = 0$; the point $(1, 1)$ is not, since $1+1-1 = 1 > 0$, so it lies in the half-space where $\underset{\sim}{w} \cdot \underset{\sim}{x} + b > 0$.
- With $\underset{\sim}{w} = \underset{\sim}{0}$, the equation reads $b = 0$: if $b \neq 0$ no point satisfies it, and if $b = 0$ every point does, so the set is empty or all of $\mathbb{R}^2$, not a line. That degenerate case is why the definition requires $\underset{\sim}{w} \neq \underset{\sim}{0}$.

Exercise 0.19 (Testing whether a vertical line is a hyperplane)**.** A vertical line in $\mathbb{R}^2$ is parallel to the x_2-axis. It cannot be written as $x_2 = mx_1 + k$ for any numbers m and k. Let L be the set of points $\underset{\sim}{x} \in \mathbb{R}^2$ with $x_1 = 3$.
Is L a hyperplane in $\mathbb{R}^2$?

Hint

Look for $\underset{\sim}{w} \neq \underset{\sim}{0}$ and b in Definition 0.10.

Solution

Solution 0.19. Yes. Take $\underset{\sim}{w} = (1, 0)^\top \neq \underset{\sim}{0}$ and $b = -3$. Then

$$\underset{\sim}{w} \cdot \underset{\sim}{x} + b = x_1 - 3,$$

which is 0 exactly when $x_1 = 3$. So L is the hyperplane with these $\underset{\sim}{w}$ and b.

Exercise 0.20 (Finding a normal vector of a plane in $\mathbb{R}^3$)**.** In $\mathbb{R}^3$ a hyperplane is a plane, and a normal vector points straight out of it. Let H be the hyperplane of points $\underset{\sim}{x} \in \mathbb{R}^3$ with

$$2x_1 - x_2 + 4x_3 = 7.$$

Find a normal vector of H.

Hint

Rewrite the equation in the form of Definition 0.10.

Solution

Solution 0.20. Subtracting 7 from both sides gives

$$2x_1 - x_2 + 4x_3 - 7 = 0,$$

which is $\underset{\sim}{w} \cdot \underset{\sim}{x} + b = 0$ with $\underset{\sim}{w} = (2, -1, 4)^\top$ and $b = -7$. So $\underset{\sim}{w} = (2, -1, 4)^\top$ is a normal vector of H. Any nonzero multiple of it is one too, since multiplying the equation by a nonzero number leaves H unchanged.

The course website's version of this page has an interactive plot here.

Exercise 0.21 (Write a rule in matrix form)**.** A clinic scores a patient's risk from three measurements.

symbol	meaning	units
x_1	age	years
x_2	systolic blood pressure	mmHg
x_3	body mass index	kg/m^2

The rule in use is

$$\text{score} = 0.4\,x_1 + 0.2\,x_2 + 1.5\,x_3 - 30$$

Write this rule in the form of Equation 1, naming $\underset{\sim}{w}$, b and p explicitly.

Solution

Solution 0.21. Read the coefficients straight off the rule. There are three inputs, so $p = 3$:

$$\underset{\sim}{w} = \begin{bmatrix} 0.4 \\ 0.2 \\ 1.5 \end{bmatrix}, \qquad b = -30$$

and the rule is $f(\underset{\sim}{x}) = \underset{\sim}{w}^\top \underset{\sim}{x} + b$.

Exercise 0.22 (Score a patient)**.** Give the score from the rule in Exercise 0.21 for a patient aged 50 with blood pressure 130 and BMI 28.

Solution

Solution 0.22. For the patient,

$$\underset{\sim}{x} = \begin{bmatrix} 50 \\ 130 \\ 28 \end{bmatrix}$$

$$\begin{aligned} \underset{\sim}{w}^\top \underset{\sim}{x} &= (0.4)(50) + (0.2)(130) + (1.5)(28) \\ &= 20 + 26 + 42 \\ &= 88 \end{aligned}$$

$$\begin{aligned} f(\underset{\sim}{x}) &= 88 - 30 \\ &= 58 \end{aligned}$$

Two things are worth noticing. The coefficients are not comparable as they stand. The weight on BMI, 1.5, is 7.5 times the weight on blood pressure, 0.2, but the contribution of BMI, 42, is less than twice the contribution of blood pressure, 26, because the BMI value, 28, is much smaller than the blood pressure value, 130. So the weight on a feature says nothing on its own about how much that feature contributes to the score. And $b = -30$ changes the score: without it, the score would be 88 instead of 58.

Exercise 0.23 (Which of these are affine in $\underset{\sim}{w}$?)**.** Each of the five expressions is a function of the weights $\underset{\sim}{w} \in \mathbb{R}^p$, with the data $\underset{\sim}{x}$ and $\underset{\sim}{y}$ held fixed. Say which ones are affine, that is, have the form Equation 1 — an inner product plus an offset — and which of the affine ones are also linear.

1. $\underset{\sim}{w}^\top \underset{\sim}{x}$
2. $\underset{\sim}{w}^\top \underset{\sim}{x} - y$
3. $(\underset{\sim}{w}^\top \underset{\sim}{x} - y)^2$
4. $\|\underset{\sim}{w}\|_2^2$
5. $w_1 x_1 + 3$

Solution

Solution 0.23.

1. **Affine, and linear.** It is Equation 1 with $b = 0$, which is a linear map by Theorem 0.2: the matrix $\underset{\sim}{x}^\top$ times $\underset{\sim}{w}$.
2. **Affine.** Still an inner product plus an offset; here the offset is $b = -y$, which is a constant because y is held fixed. It is linear only when $y = 0$: otherwise $\underset{\sim}{w} = \underset{\sim}{0}$ gives $-y \neq 0$, and a linear map sends $\underset{\sim}{0}$ to 0 (Theorem 0.1).
3. **Not affine.** Squaring is not an inner product plus an offset. This expression is the squared error of one observation, and its curvature in $\underset{\sim}{w}$ is exactly why fitting a model is an optimization problem rather than a lookup.
4. **Not affine.** By Equation 5 in Vectors, the squared norm $\|\underset{\sim}{w}\|_2^2$ is $\underset{\sim}{w}^\top \underset{\sim}{w}$, which has $\underset{\sim}{w}$ in both slots at once. Hold either slot fixed and it is linear in the other; as a function of the single variable $\underset{\sim}{w}$ it is quadratic.
5. **Affine, not linear.** It is Equation 1 with coefficient vector $\underset{\sim}{a} = (x_1, 0, \ldots, 0)^\top$ and $b = 3$; at $\underset{\sim}{w} = \underset{\sim}{0}$ it equals $3 \neq 0$, so it is not linear (Theorem 0.1). The data $\underset{\sim}{x}$ is untouched — $\underset{\sim}{a}$ is a new vector built from its first entry.

Items 1, 2 and 5 are affine and items 3 and 4 are not, which is the split that matters: a *model* can be affine in its weights while the *quantity we minimize* is not. Squared error and the norm penalty are both curved in $\underset{\sim}{w}$, and that curvature is what makes fitting an optimization problem at all: there is always a downhill direction to follow. Curvature alone does not promise a *single* answer. The squared error of one observation is completely flat along a whole line of weight vectors, every one of which fits that observation exactly, and pinning down one of them is part of what the norm penalty is added for in week 3.

[1] Wikipedia contributors, "Linear map — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Linear_map

[2] Wikipedia contributors, "Translation (geometry) — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Translation_(geometry)

[3] H. Zhou, "Matrices." Lecture notes for Biostat 216, Mathematical Methods for Biostatistics, University of California, Los Angeles, 2024. Accessed: Oct. 05, 2026. [Online]. Available: https://ucla-biostat-216.github.io/2024fall/slides/03-matrix/03-matrix.html

[4] S. Boyd and L. Vandenberghe, *Introduction to applied linear algebra: Vectors, matrices, and least squares*. Cambridge University Press, 2018. doi: 10.1017/9781108583664.

[5] Wikipedia contributors, "Hyperplane — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Hyperplane

[6] Wikipedia contributors, "Normal (geometry) — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Normal_(geometry)