Example of linear functions and how to write them as matrices-vector products

**Example 4.1.** The function $f_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ that rotates a vector $x$ by $\theta$ counterclockwise is linear and can be written as

$$f(x) = \begin{pmatrix} \cos \theta & x_1 - \sin \theta & x_2 \\ \sin \theta & x_1 + \cos \theta & x_2 \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Hence

$$= : R_\theta \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (4.1)$$

A matrix of the form $R_\theta$ is called a rotation matrix.

Using $\sin \frac{\pi}{2} = 1$ and $\cos \frac{\pi}{2} = 0$ we can check that the first matrix of Example 3.11 is indeed a rotation matrix with $\theta = \frac{\pi}{2}$ (i.e. 90 deg).

**Example 4.2 (Population Dynamics).** We study the population size over multiple generations of spotted owls, where we distinguish between 3 stages: young, adult and dead.

From one generation to the next, the following happens:

- 30% of the young reach adulthood, 70% die.
- 50% of the adults survive the winter, 50% die.
- on average, each adult has 1.3 young as offspring.

We call the current state: $x = \begin{pmatrix} x_{\text{young}} \\ x_{\text{adult}} \end{pmatrix}$ and the future state $y = \begin{pmatrix} y_{\text{young}} \\ y_{\text{adult}} \\ y_{\text{dead}} \end{pmatrix}$. 

23
Then we can describe the population dynamic by a linear function
\[
\begin{pmatrix}
  y_{\text{young}} \\
  y_{\text{adult}} \\
  y_{\text{dead}}
\end{pmatrix} = \begin{pmatrix}
  1.3 & x_{\text{adult}} \\
  0.3 \cdot x_{\text{young}} + 0.5 \cdot x_{\text{adult}} \\
  0.7 \cdot x_{\text{young}} + 0.5 \cdot x_{\text{adult}}
\end{pmatrix} = \begin{pmatrix}
  0 & 1.3 \\
  0.5 & 0 \\
  0.7 & 0.5
\end{pmatrix} \begin{pmatrix}
  x_{\text{young}} \\
  x_{\text{adult}}
\end{pmatrix} \in \mathbb{R}^3
\] (4.2)

**Definition 4.3.** For any polynomial 
\[ p(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0, \]
we call 
\[ x_p := \begin{pmatrix}
  a_n \\
  a_{n-1} \\
  \vdots \\
  a_1 \\
  a_0
\end{pmatrix} \in \mathbb{R}^{n+1} \] its coefficient vector.

**Example 4.4 (Derivatives of polynomials).** The function \( D : \mathbb{R}^{n+1} \to \mathbb{R}^n \) that maps the coefficient vector of a polynomial to the coefficient vector of its derivative is a linear function.

Check: The derivative of any polynomial
\[
p(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0
\] (4.3)
is
\[
p'(t) = n a_n t^{n-1} + (n - 1) a_{n-1} t^{n-2} + \cdots + 2 a_2 t + a_1,
\] (4.4)
which has the coefficient vector 
\[ x_{p'} = \begin{pmatrix}
  n a_n \\
  (n - 1) a_{n-1} \\
  \vdots \\
  2 a_2 \\
  a_1
\end{pmatrix} \in \mathbb{R}^n. \]

We be write \( D \) as
Example 4.5 (Slopes of discretized functions).

Let $x^{(1)}, x^{(n)}$ be $n$ increasing positions on the $x$-axis, and let $y^{(1)}, \ldots, y^{(n)}$ be the values of a function at this position. We can form the piece-wise linear (actually: piece-wise affine) approximation to the unknown function by "connecting the dots" (see Figure).

For $i = 1, \ldots, n - 1$, we can estimate the slope of the piecewise linear approximation by $s_i = \frac{dy}{dx} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}$. The function $D$ that maps the values
LECTURE 4. EXAMPLES OF MATRIX OPERATIONS

$y_1, \ldots, y_n$ to the slopes $s_1, \ldots, s_{n-1}$ is linear:

$$D(y) = \begin{pmatrix} s_1 \\ \vdots \\ s_{n-1} \end{pmatrix} = \begin{pmatrix} \frac{y_2 - y_1}{x_2 - x_1} \\ \vdots \\ \frac{y_n - y_{n-1}}{x_n - x_{n-1}} \end{pmatrix}$$ (4.9)

$$= \begin{pmatrix} \frac{-1}{x_2-x_1} & \frac{1}{x_2-x_1} & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{x_3-x_2} & \frac{1}{x_3-x_2} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \frac{-1}{x_n-x_{n-1}} & \frac{1}{x_n-x_{n-1}} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix}$$ (4.10)

**Remark.** The rule how to multiply a matrix $A \in \mathbb{R}^{m \times n}$ with a vector $x \in \mathbb{R}^n$ is typically written as

$$[Ax]_i = \sum_{j=1}^{n} a_{ij} x_j$$ (matrix–vector multiplication) (4.11)

for $i = 1, \ldots, m$. ($Ax$ is an $m$-vector, and this explains how to compute its $m$ components).

How to remember?

1) Each of the $m$ rows in matrix defines linear function on input, result of that goes into corresponding output row

$$\begin{pmatrix} * & * & * & * \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} * \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} = \begin{pmatrix} * \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}$$

each row/column pair: multiply componentwise and add up

2) Alternative view: Think of $A$ as collection of $n$ columns: $(c^{(1)}|c^{(2)}|\ldots|c^{(n)})$, so for each $j = 1, \ldots, n$, $c^{(j)} \in \mathbb{R}^m$ and $[c^{(j)}]_i = a_{ij}$.

$$[Ax]_i = \sum_{j=1}^{n} a_{ij} x_j = \sum_{j=1}^{n} x_j \left[c^{(j)}\right]_i = \left[\sum_{j=1}^{n} x_j c^{(j)}\right]_i$$ (4.12)

$$Ax = \sum_{j=1}^{n} x_j c^{(j)}$$ (4.13)

The output vector is a weighted sum (called "linear combination") of the matrix columns with coefficients given by the entries of $x$. 
Note: Both views are equivalent, use whichever is more convenient. I use \( ii \) if \( x \) has very few non-zero entries, otherwise \( i \).

Also note: Both versions only make sense if the number of columns of the matrix is the same as the number of entries (rows) of the vector.

Example 4.6.

\[
\begin{pmatrix}
1 & 2 & 3 \\
3 & 2 & 1
\end{pmatrix}
\begin{pmatrix}
\frac{1}{3} \\
\frac{2}{3}
\end{pmatrix} = \left(1 \cdot \frac{1}{3} + 2 \cdot \frac{2}{3} + 3 \cdot \frac{3}{3}\right) = \left(\frac{1}{3} + 4 + 9\right) = \left(\frac{14}{3} + 4 + 3\right) = \left(\frac{14}{10}\right)
\]

(4.14)

Example 4.7.

\[
\begin{pmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
7 & 8 & 9
\end{pmatrix}
\begin{pmatrix}
0 \\
-1 \\
0
\end{pmatrix} = \left(0 + 2 \cdot (-1) + 0\right) = \left(-2\right)
\]

(4.16)

\[
\begin{pmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
7 & 8 & 9
\end{pmatrix}
\begin{pmatrix}
0 \\
-1 \\
0
\end{pmatrix} = -1 \begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix} = \begin{pmatrix} -2 \\ -5 \\ -8 \end{pmatrix}
\]

(4.17)