This module:
will acquaint you with some basic linear algebra used later.

Suggested reading:
Part of this relates to Section 5.5 (5.4 in 3rd edition).
But this material is mostly covered in texts on linear algebra.

Chapter 5: Random Vectors

Motivation
In this chapter we go beyond 1 or 2 random variables to $n$.
The key issue is that $n$ is large enough to make it difficult to model the pdf. But more on that later...

Why do we even need to consider $n > 2$ random variables over the same probability space?
Example application

Seismic discrimination — Given samples of a waveform over time \( X(t), \) \( t \in \{t_1, t_2, \ldots, t_n\} \), we want to determine whether the waveform was generated by an earthquake or an underground explosion.

We will compare the vector \( X \) to stored data and decide which type of event generated it.

Example application

Disease detection — We take multiple measurements of a patient (in the book they mention looking for black-lung disease), and want to decide whether the patient has the disease.

Challenge (p302 in 4th edition)

If there were \( n = 2 \) or \( 3 \) random variables, we could understand the distribution and process data manually.

But when \( n \) is large, the different elements \( X_i \) are dependent, and we have no convenient features (such as Gaussianity), even estimating the distribution is tough.
To approach these types of problems, a convenient engineering design methodology considers the two main moments of the data: the expected value and the covariances. To work with these, we will begin with a review of linear algebra.

**Linear algebra review**

**Matrix notation**

\[
A = \begin{bmatrix}
1 & 2 \\
3 & 4 \\
5 & 6
\end{bmatrix}
= [a_{11}, a_{12}, \\
a_{21}, a_{22}, \\
a_{31}, a_{32}]
\]

A is of size $3 \times 2$.
The first subscript is for the row, the second for the column.

**Transpose**

\[
A^T = \begin{bmatrix}
1 & 3 & 5 \\
2 & 4 & 6
\end{bmatrix}
= [a_{11}, a_{21}, a_{31}, \\
a_{12}, a_{22}, a_{32}]
\]

**Symmetric matrix**

\[A = A^T\]

**Matrix addition/subtraction**

\[A_{mxn} + B_{mxn} = C_{mxn}\]
\[a_{ij} + b_{ij} = c_{ij}\]
**Matrix product**  
\[ A^{m \times k} \times B^{k \times n} = C^{m \times n} \]  
\[ c_{ij} = \sum_{k=1}^{k} a_{ik}b_{kj} \]

**Example**  
\[
\begin{bmatrix}
1 & 2 \\
3 & 4
\end{bmatrix}
\begin{bmatrix}
1 & 2 \\
3 & 4
\end{bmatrix}
= \begin{bmatrix}
1\cdot1 + 2\cdot3 & 1\cdot2 + 2\cdot4 \\
3\cdot1 + 4\cdot3 & 3\cdot2 + 4\cdot4
\end{bmatrix}
= \begin{bmatrix}
7 & 10 \\
15 & 22
\end{bmatrix}
\]

**Column vector**  \[ m \times 1 \text{ matrix} \]  
**Row vector**  \[ 1 \times m \text{ matrix} \]

**Euclidean length**  or  **norm**  
\[ x = (x_1, \ldots, x_n) \]  
\[ \|x\|_2 = \left(\sum_{i=1}^{n} (x_i)^2\right)^{\frac{1}{2}} \]

**p-norm**  
\[ \|x\|_p = \left(\sum_{i=1}^{n} (x_i)^p\right)^{\frac{1}{p}} \]

**Example**  
\[ x = \begin{bmatrix}
1 \\
2
\end{bmatrix}\]

\[ \|x\|_1 = \sqrt{1^2 + 2^2} = \sqrt{5} \]
\[ \|x\|_3 = \sqrt[3]{1^3 + 2^3} = 3\sqrt[3]{9} \]
Applications

Where does this appear?

Suppose we measure temperature once an hour for a day. 

\[ X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{24} \end{pmatrix} \]

A digital camera captures 1,000 x 4,000 pixel images.

\[ A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,1000} \\ a_{21} & a_{22} & \cdots & a_{2,1000} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1000,1} & a_{1000,2} & \cdots & a_{1000,1000} \end{bmatrix} \]

**Eigen-values**

\[ A \cdot v = \lambda \cdot v \]

\[ \uparrow \quad \uparrow \]

**eigen-vector**  **scalar**  **e-value**

**Example**

\[ A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \]

\[ \det \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} = (2-\lambda)^2 - 1 = 0 \]

\[ (3-\lambda)(1-\lambda) = 0 \]

\[ \Rightarrow \lambda_1 = 3, \quad \lambda_2 = 1 \]
\( (A-\lambda I) V = (0) \)

\( \lambda_1 = 3 \quad A-\lambda I = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \)

\[ \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \]

\[ \Rightarrow \quad v_1 = v_2 \]

In the book, they require unit norm, \( v_1^2 + v_2^2 = 1 \)

\[ \Rightarrow \quad v_1 = v_2 = \frac{1}{\sqrt{2}}, \quad v = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \]

\( \lambda_2 = 1 \quad A-\lambda I = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \)

\[ \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \]

\[ v_1^2 + v_2^2 = 1 \]

\[ \Rightarrow \quad v_1 = \frac{1}{\sqrt{2}}, \quad v_2 = -\frac{1}{\sqrt{2}}, \quad v = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \]
Determinant analogous to amplification of volume.

\[ \text{Vol}(A_s) = |\det(A)| \cdot \text{Vol}(S) \]

In the 2x2 case,

\[
\det(A) = \det \begin{pmatrix}
    a_{11} & a_{12} \\
    a_{21} & a_{22}
\end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}
\]

In the general case, we have for diagonal matrices

\[
\det \begin{pmatrix}
    a_1 & 0 & \cdots & 0 \\
    0 & a_2 & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & a_n
\end{pmatrix} = a_1a_2\cdots a_n
\]