LECTURE NOTES[1]
AN INTRODUCTION TO DIGITAL COMMUNICATIONS

Armand M. Makowski [2]

[1] ©1997-2011 by Armand M. Makowski

[2] Department of Electrical and Computer Engineering, and Institute for Systems Research, University of Maryland, College Park, MD 20742. E-mail: armand@isr.umd.edu. Phone: (301) 405-6844

Part I

Preliminaries

Chapter 1

Decision Theory

In this chapter we present the basic ideas of statistical decision theory that will be used repeatedly in designing optimal receivers in a number of settings. These design problems can all be reduced to problems of M-ary hypothesis testing which we investigate below in generic form.

1.1 The generic hypothesis testing problem

In the statistical hypothesis testing problem, a decision has to be made as to which of several possible hypotheses (or states of nature) is the correct one. The state of nature is encoded in a rv H and a decision has to be made on the basis of an $\mathbb{R}^d$-valued observation vector $\boldsymbol{X}$ which is statistically related to H. Given that a cost is incurred for making decisions, the decision-maker seeks to determine the "best" decision to be implemented. Although several formulations are available in the literature, here we concentrate on the Bayesian formulation.

1.1.1 The Bayesian model

Let $\mathcal{H}$ denote a *finite* set with M elements for some positive integer $M \geq 2$, say $\mathcal{H} := \{1, \ldots, M\}$ for the sake of concreteness. The rv H takes values in $\mathcal{H}$ according to the pmf

$$p_m := \mathbb{P}\left[H = m\right], \quad m = 1, \ldots, M.$$

This pmf $\boldsymbol{p} = (p_1, \ldots, p_M)$ is often called the *prior* on H.

With each of the possible hypothesis $m = 1, \ldots, M$, we associate a probability distribution function F_m on $\mathbb{R}^d$ with the interpretation that F_m is the conditional distribution of $\boldsymbol{X}$ given $H = m$, i.e.,

$$\mathbb{P}\left[\boldsymbol{X} \leq \boldsymbol{x} | H = m\right] = F_m(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{R}^d.$$

The observation rv $\boldsymbol{X}$ is then distributed according to

$$\mathbb{P}\left[\boldsymbol{X} \leq \boldsymbol{x}\right] = \sum_{m=1}^{M} p_m F_m(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{R}^d$$

by the Law of Total Probabilities, while

$$\mathbb{P}\left[\boldsymbol{X} \leq \boldsymbol{x}, H = m\right] = p_m F_m(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{R}^d, \ m = 1, \ldots, M.$$

In other words, the conditional probability distribution of the observations given the hypothesis *and* the probability distribution of H completely specify the *joint* distribution of the rvs H and $\boldsymbol{X}$.

1.1.2 The optimization problem

On observing the observation vector, the decision-maker implements a decision rule which returns a state of nature in response to this observation. Thus, an (admissible) decision rule or *detector*[1] is simply any mapping $d : \mathbb{R}^d \to \mathcal{H}$.[2] In the language of Estimation Theory, the mapping $\delta : \mathbb{R}^d \to \mathcal{H}$ can be interpreted as an *estimator* for H (on the basis of $\boldsymbol{X}$) with $\delta(\boldsymbol{X})$ representing the corresponding *estimate* $\widehat{H}$ of H (on the basis of $\boldsymbol{X}$). Let $\mathcal{D}$ denote the class of all (admissible) detection rules.

As a cost is incurred for making decisions, we introduce the mapping $C : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ with the interpretation that

$$C(m, k) = \begin{array}{c}\text{Cost incurred for deciding } k \\ \text{when } H = m\end{array}$$

[1] In the statistical literature on Hypothesis Testing such a detector is often called a *test*, while in the context of Digital Communications, a detector is often refered to as a *receiver* for reasons that will become shortly apparent – We shall follow this tradition in due time!

[2] Strictly speaking, the definition of an admissible rule should include the property that each of the sets $\{\boldsymbol{x} \in \mathbb{R}^d : \ \delta(\boldsymbol{x}) = m\}$, $m = 1, \ldots, M$, be a Borel subset of $\mathbb{R}^d$.

for all $k, m = 1, \ldots, M$. The use of any admissible rule d in $\mathcal{D}$ thus incurs a cost $C(H, \delta(\boldsymbol{X}))$. However, the value of the cost $C(H, \delta(\boldsymbol{X}))$ is not available to the decision-maker[3] and attention focuses instead on the *expected cost* $J : \mathcal{D} \rightarrow \mathbb{R}$ defined by

$$J(\delta) := \mathbb{E}\left[C(H, \delta(\boldsymbol{X})\right], \quad \delta \in \mathcal{D}.$$

The Bayesian M-ary hypothesis testing problem (P_B) is now formulated as

$$(P_B): \quad \text{Minimize } J \text{ over the collection } \mathcal{D} \text{ of admissible decision rules}$$

Solving problem (P_B) amounts to identifying detector(s) $\delta^\star : \mathbb{R}^d \rightarrow \mathcal{H}$ such that

$$J(\delta^\star) \leq J(\delta), \quad \delta \in \mathcal{D}.$$

Any detector $\delta^\star : \mathbb{R}^d \rightarrow \mathcal{H}$ which minimizes the expected cost is referred to as an *optimal* detector.

The problem (P_B) can be solved for arbitrary cost functions C under fairly weak assumptions on the distributions $F_1, \ldots, F_M$. Throughout, to simplify matters somewhat, we assume that for each $m = 1, \ldots, M$, the distribution function F_m admits a density f_m on $\mathbb{R}^d$, i.e.,

$$F_m(\boldsymbol{x}) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_d} f_m(\boldsymbol{t}) dt_1 \ldots dt_d, \quad \boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d.$$

This assumption is enforced in all cases considered here.

Rather than discussing the case of a general cost function, we will instead focus on a special case of paramount importance to Digital Communications. This occurs when C takes the form

$$C(m, k) = \begin{cases} 1 & \text{if } m \neq k \\ \\ 0 & \text{if } m = k \end{cases}, \quad k, m = 1, \ldots, M \tag{1.1}$$

and the expected cost reduces to the so-called *probability of error*

$$\mathtt{Er}(\delta) := \mathbb{P}\left[\delta(\boldsymbol{X}) \neq H\right], \quad \delta \in \mathcal{D}. \tag{1.2}$$

Versions of the problem with cost (1.1)–(1.2) will be extensively discussed in this text. The remainder of the discussion assumes this cost structure.

[3] Indeed the value of H is not known, in fact needs to be estimated!

1.2 Identifying the optimal detector

As the first step in solving the problem (P_B), we argue now as to the form of the optimal detector. We begin by noting that any detector $\delta : \mathbb{R}^d \rightarrow \mathcal{H}$ is equivalent to a *partition* $(\Delta_1, \ldots, \Delta_M)$ of $\mathbb{R}^d$, that is, a collection of subsets of $\mathbb{R}^d$ such that

$$\Delta_m \cap \Delta_k = \emptyset, \qquad \begin{array}{c} k \neq m \\ k, m = 1, \ldots, M \end{array}$$

with

$$\mathbb{R}^d = \cup_{m=1}^M \Delta_m.$$

Indeed, any detector $\delta : \mathbb{R}^d \rightarrow \mathcal{H}$ induces a partition $(\Delta_1, \ldots, \Delta_M)$ of $\mathbb{R}^d$ by setting

$$\Delta_m = \{\boldsymbol{x} \in \mathbb{R}^d : \ \delta(\boldsymbol{x}) = m\}, \quad m = 1, \ldots, M.$$

Conversely, with any partition $(\Delta_1, \ldots, \Delta_M)$ of $\mathbb{R}^d$ we can associate a detector $d : \mathbb{R}^d \rightarrow \mathcal{H}$ through the correspondence

$$d(\boldsymbol{x}) = m \quad \text{if} \quad \boldsymbol{x} \in \Delta_m, \quad m = 1, \ldots, M.$$

Start with a detector $\delta : \mathbb{R}^d \rightarrow \mathcal{H}$ with induced partition $(\Delta_1, \ldots, \Delta_M)$ as above. We have

$$\begin{aligned} \mathbb{P}\left[\delta(\boldsymbol{X}) = H\right] &= \sum_{m=1}^M p_m \mathbb{P}\left[\delta(\boldsymbol{X}) = m | H = m\right] \\ &= \sum_{m=1}^M p_m \mathbb{P}\left[\boldsymbol{X} \in \Delta_m | H = m\right] \\ &= \sum_{m=1}^M p_m \int_{\Delta_m} f_m(\boldsymbol{x}) d\boldsymbol{x}. \end{aligned}$$

As we seek to minimize the probability of error, we conclude that it suffices to maximize

$$\begin{aligned} F(\Delta_1, \ldots, \Delta_M) &:= \sum_{m=1}^M p_m \int_{\Delta_m} f_m(\boldsymbol{x}) d\boldsymbol{x} \\ &= \int_{\mathbb{R}^d} \left(\sum_{m=1}^M \mathbf{1}\left[\boldsymbol{x} \in \Delta_m\right] p_m f_m(\boldsymbol{x}) \right) d\boldsymbol{x} \end{aligned}$$

with respect to partitions $(\Delta_1, \ldots, \Delta_M)$ of $\mathbb{R}^d$.

Inspection of the functional F suggests a possible candidate for optimality: For each $m = 1, \ldots, M$, set

$$\Delta^\star_m := \{\boldsymbol{x} \in \mathbb{R}^d : \ p_m f_m(\boldsymbol{x}) = \max_{k=1,\ldots,M} p_k f_k(\boldsymbol{x})\}$$

with tie breakers if necessary. For sake of concreteness, ties are broken according to the lexicographic order, i.e., if at point $\boldsymbol{x}$, it holds that

$$p_i f_i(\boldsymbol{x}) = \max_{k=1,\ldots,M} p_k f_k(\boldsymbol{x}) = p_j f_j(\boldsymbol{x})$$

for distinct values i and j, then $\boldsymbol{x}$ will be assigned to $\Delta^\star_i$ if $i < j$. With such precautions, these sets form a partition $(\Delta^\star_1, \ldots, \Delta^\star_M)$ of $\mathbb{R}^d$, and the detector $\delta^\star : \mathbb{R}^d \to \mathcal{H}$ associated with this partition takes the form

$$\delta^\star(\boldsymbol{x}) = m \quad \text{iff} \quad p_m f_m(\boldsymbol{x}) = \max_{k=1,\ldots,M} p_k f_k(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{R}^d \tag{1.3}$$

with a lexicographic tie-breaker, or more compactly,

$$\delta^\star(\boldsymbol{x}) = \arg\max\left(m = 1, \ldots, M : \ p_m f_m(\boldsymbol{x})\right), \quad \boldsymbol{x} \in \mathbb{R}^d.$$

We shall often write that $\delta^\star$ prescribes

$$\widehat{H} = m \quad \text{iff} \quad p_m f_m(\boldsymbol{x}) \text{ largest} \tag{1.4}$$

with the interpretation that upon collecting the observation vector $\boldsymbol{x}$, the detector $\delta^\star$ selects the state of nature m as its estimate on the basis of $\boldsymbol{x}$.

1.3 The detector $\delta^\star$ is optimal

That the guess (1.4) is indeed correct forms the content of the next proposition:

Theorem 1.3.1 *The detector $\delta^\star : \mathbb{R}^d \to \mathcal{H}$ given by (1.3) is optimal, in that $\mathtt{Er}(\delta^\star) \leq \mathtt{Er}(\delta)$ for any other detector $\delta : \mathbb{R}^d \to \mathcal{H}$.*

Proof. Introduce the mapping $f : \mathbb{R}^d \to \mathbb{R}$ by

$$f(\boldsymbol{x}) = \max_{m=1,\ldots,M} p_m f_m(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{R}^d.$$

The obvious bound

$$f(\boldsymbol{x}) \leq \sum_{m=1}^{M} p_m f_m(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{R}^d$$

implies

$$\int_{\mathbb{R}^d} f(\boldsymbol{x}) d\boldsymbol{x} \leq \sum_{m=1}^{M} p_m \int_{\mathbb{R}^d} f_m(\boldsymbol{x}) d\boldsymbol{x} = \sum_{m=1}^{M} p_m = 1,$$

and the function f is indeed integrable over all of $\mathbb{R}^d$. This fact will be used without further mention in the discussion below to validate some of the manipulations involving integrals.

For any partition $(\Delta_1, \ldots, \Delta_M)$ of $\mathbb{R}^d$, we need to show that

$$F(\Delta_1^\star, \ldots, \Delta_M^\star) - F(\Delta_1, \ldots, \Delta_M) \geq 0, \tag{1.5}$$

where

$$\begin{aligned} & F(\Delta_1^\star, \ldots, \Delta_M^\star) - F(\Delta_1, \ldots, \Delta_M) \\ = \ & \sum_{m=1}^{M} \left(\int_{\Delta_m^\star} p_m f_m(\boldsymbol{x}) d\boldsymbol{x} - \int_{\Delta_m} p_m f_m(\boldsymbol{x}) d\boldsymbol{x} \right). \end{aligned}$$

Next, for each $m = 1, \ldots, M$, by the definition of $\Delta_m^\star$ and f it holds that

$$p_m f_m(\boldsymbol{x}) = f(\boldsymbol{x}), \quad \boldsymbol{x} \in \Delta_m^\star$$

and

$$p_m f_m(\boldsymbol{x}) \leq f(\boldsymbol{x}), \quad \boldsymbol{x} \in \Delta_m.$$

Therefore,

$$\begin{aligned} & F(\Delta_1^\star, \ldots, \Delta_M^\star) - F(\Delta_1, \ldots, \Delta_M) \\ = \ & \sum_{m=1}^{M} \left(\int_{\Delta_m^\star} f(\boldsymbol{x}) - \int_{\Delta_m} p_m f_m(\boldsymbol{x}) d\boldsymbol{x} \right) \\ \geq \ & \sum_{m=1}^{M} \left(\int_{\Delta_m^\star} f(\boldsymbol{x}) d\boldsymbol{x} - \int_{\Delta_m} f(\boldsymbol{x}) d\boldsymbol{x} \right) \\ = \ & \sum_{m=1}^{M} \int_{\Delta_m^\star} f(\boldsymbol{x}) d\boldsymbol{x} - \sum_{m=1}^{M} \int_{\Delta_m} f(\boldsymbol{x}) d\boldsymbol{x} \\ = \ & \int_{\mathbb{R}^d} f(\boldsymbol{x}) d\boldsymbol{x} - \int_{\mathbb{R}^d} f(\boldsymbol{x}) d\boldsymbol{x} = 0, \end{aligned}$$

and the inequality (1.5) is established. ■

1.4 Alternate forms of the optimal detector

The optimal detector $\delta^\star$ identified in Theorem 1.3.1 is amenable to useful interpretations which we now develop

The MAP detector With the usual caveat on tie breakers, the definition (1.3) of the optimal detector $\delta^\star$ yields

$$\begin{aligned}
\text{Choose } \widehat{H} = m \quad & \text{iff} \quad p_m f_m(\boldsymbol{x}) \text{ largest} \\
& \text{iff} \quad \frac{p_m f_m(\boldsymbol{x})}{\sum_{k=1}^M p_k f_k(\boldsymbol{x})} \text{ largest} \\
& \text{iff} \quad \mathbb{P}\left[H = m | \boldsymbol{X} = \boldsymbol{x}\right] \text{ largest}
\end{aligned}$$

where the last equivalence follows from Bayes' Theorem in the form

$$\mathbb{P}\left[H = m | \boldsymbol{X} = \boldsymbol{x}\right] = \frac{p_m f_m(\boldsymbol{x})}{\sum_{k=1}^M p_k f_k(\boldsymbol{x})}, \quad \boldsymbol{x} \in \mathbb{R}^d$$

for each $m = 1, \ldots, M$. In particular, $\delta^\star$ can be viewed as selecting $\widehat{H} = m$ whenever the *a posteriori* probability of H given the "observations" $\boldsymbol{X}$ is largest. In the parlance of Estimation Theory, $\delta^\star$ is the *Maximum A Posteriori* (MAP) estimator of the "parameter" H on the basis of the observations $\boldsymbol{X}$.

As monotone increasing transformations are order preserving, the optimal detector $\delta^\star$ has the equivalent form

$$\text{Choose } \widehat{H} = m \quad \text{iff} \quad \log\left(p_m f(\boldsymbol{x}|H = m)\right) \text{ largest.}$$

Uniform prior and the ML detector There is one situation of great interest, from both practical and theoretical viewpoints, where further simplifications are achieved in the structure of the optimal detector. This occurs when the rv H is *uniformly* distributed over $\mathcal{H}$, namely

$$\mathbb{P}\left[H = m\right] = \frac{1}{M}, \quad m = 1, \ldots, M. \qquad (1.6)$$

In that case, the optimal detector $\delta^\star$ prescribes

$$\text{Choose } \widehat{H} = m \quad \text{iff} \quad f_m(\boldsymbol{x}) \text{ largest,}$$

and therefore implements the so-called *Maximum Likelihood* (ML) estimate of H on the basis of $\boldsymbol{x}$.

1.5 An important example

An important special case arises when the distributions $F_1, \ldots, F_M$ are all Gaussian distributions with the same *invertible* covariance matrix. This is equivalent to

$$[\boldsymbol{X}|H = m] =_{st} \boldsymbol{\mu}_m + \boldsymbol{V}, \quad m = 1, \ldots, M \tag{1.7}$$

where $\boldsymbol{V}$ is a zero mean $\mathbb{R}^d$-valued Gaussian rv with covariance matrix $\boldsymbol{\Sigma}$. We assume $\boldsymbol{\Sigma}$ to be invertible and the mean vectors $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_M$ to be distinct. An alternative description, based on (1.7), relates the observation $\boldsymbol{X}$ to the state of nature H through the measurement equation

$$\boldsymbol{X} = \boldsymbol{\mu}_H + \boldsymbol{V} \tag{1.8}$$

where the rvs H and $\boldsymbol{V}$ are assumed to be mutually independent rvs distributed as before. Under this observation model, for each $m = 1, \ldots, M$, F_m admits the density

$$f_m(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^p \det(\boldsymbol{\Sigma})}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_m)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_m)}, \quad \boldsymbol{x} \in \mathbb{R}^d. \tag{1.9}$$

We note that

$$\begin{aligned} &\log\left(p_m f_m(\boldsymbol{x})\right) \\ &\quad = \quad C + \log p_m - \frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu}_m)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}_m), \qquad \begin{array}{c} \boldsymbol{x} \in \mathbb{R}^d, \\ m = 1, \ldots, M \end{array} \end{aligned} \tag{1.10}$$

with constant C given by

$$C := -\frac{1}{2} \log\left((2\pi)^d \det(\boldsymbol{\Sigma})\right).$$

This constant being independent of m and $\boldsymbol{x}$, the optimal detector prescribes

$$\text{Choose } \widehat{H} = m \quad \text{iff} \quad 2\log p_m - (\boldsymbol{x} - \boldsymbol{\mu}_m)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}_m) \text{ largest.}$$

Under uniform prior, this MAP detector becomes the ML detector and takes the form

$$\text{Choose } \widehat{H} = m \quad \text{iff} \quad (\boldsymbol{x} - \boldsymbol{\mu}_m)' \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}_m) \text{ smallest.}$$

The form of the MAP detector given above very crisply illustrates how the prior information (p_m) on the hypothesis is modified by the posterior information collected through the observation vector $\boldsymbol{x}$. Indeed, at first, if only the prior distribution were known, and with no further information available, it is reasonable to select the most likely state of nature $H = m$, i.e., the one with largest value of p_m. However, as the observation vector $\boldsymbol{x}$ becomes available, its closeness to $\boldsymbol{\mu}_m$ should provide some indication on the underlying state of nature. More precisely, if $\boldsymbol{\mu}_m$ is the "closest" (in some sense) to the observation $\boldsymbol{x}$ among all the vectors $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_M$, then this should be taken as an indication of high likelihood that $H = m$; here the appropriate notion of closeness is the norm on $\mathbb{R}^d$ induced by $\boldsymbol{\Sigma}^{-1}$. The MAP detector combines these two trends when constructing the optimal decision in the following way: The state of nature $H = m$ may have a rather small value for its prior p_m, making it *a priori* unlikely to be the underlying state of nature, yet this will be offset if the observation $\boldsymbol{x}$ yields an extremely small value for the "distance" $(\boldsymbol{x} - \boldsymbol{\mu}_m)' \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}_m)$ to the mean vector $\boldsymbol{\mu}_m$.

When $\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{I}_d$ for some $\sigma > 0$, the components of $\boldsymbol{V}$ are mutually independent, and the MAP and ML detectors take the simpler forms

$$\text{Choose } \widehat{H} = m \quad \text{iff} \quad 2 \log p_m - \frac{1}{\sigma^2} \|\boldsymbol{x} - \boldsymbol{\mu}_m\|^2 \text{ largest}$$

and

$$\text{Choose } \widehat{H} = m \quad \text{iff} \quad \|\boldsymbol{x} - \boldsymbol{\mu}_m\|^2 \text{ smallest,}$$

respectively. Thus, given the observation vector $\boldsymbol{x}$, the ML detector returns the state of nature m whose mean vector $\boldsymbol{\mu}_m$ is closest (in the usual Euclidean sense) to $\boldsymbol{x}$. This is an example of *nearest-neighbor* detection.

1.6 Consecutive observations

As the discussion in Section 1.5 already shows, the MAP and ML detectors can assume simpler forms in structured situations. In the present section we explore possible simplifications when *repeated* observations of the state of nature are made.

A convenient setup to carry out the discussion is as follows: Consecutive observations are collected at time epochs labelled $i = 1, \ldots, n$ with $n > 1$. At each time epoch, nature is assumed to be in one of L distinct states, labelled $\ell = 1, \ldots, L$, and we write $\mathcal{L} = \{1, \ldots, L\}$. For each $i = 1, \ldots, n$, the unknown state of nature at epoch i is encoded in the $\mathcal{L}$-valued rv H_i, while the observation is modeled by an $\mathbb{R}^d$-valued rv $\boldsymbol{X}_i$. The "global" state of nature over these n time epochs is the $\mathcal{L}^n$-valued rv $\boldsymbol{H} = (H_1, \ldots, H_n)$, while the $\mathbb{R}^{nd}$-valued rv $\boldsymbol{X} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$ represents the cumulative observation over these same epochs.

The problem of interest here is that of detecting the global state of nature $\boldsymbol{H}$ on the basis of the cumulative observation vector $\boldsymbol{X}$. A number of assumptions will now be made; they are present in some situations relevant to Digital Communications: At this point, the $\mathcal{L}^n$-valued rv $\boldsymbol{H}$ is assumed to have an arbitrary pmf, say

$$\begin{aligned} p(\boldsymbol{h}) &= \mathbb{P}\left[\boldsymbol{H} = \boldsymbol{h}\right] \\ &= \mathbb{P}\left[H_1 = h_1, \ldots, H_n = h_n\right], \quad \boldsymbol{h} = (h_1, \ldots, h_n) \in \mathcal{L}^n. \end{aligned}$$

We also assume that the observations $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ are *conditionally independent* given the global state of nature, with a conditional density of the product form

$$f_{\boldsymbol{h}}(\boldsymbol{x}) = \prod_{i=1}^{n} f_{h_i}(\boldsymbol{x}_i), \qquad \begin{array}{c} \boldsymbol{h} = (h_1, \ldots, h_n) \in \mathcal{L}^n \\ \boldsymbol{x} \in \mathbb{R}^n \end{array} . \tag{1.11}$$

Note that the functional form of (1.11) implies more than the conditional independence of the rvs $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ as it also stipulates for each $i = 1, \ldots, n$ that the conditional distribution of $\boldsymbol{X}_i$ given $\boldsymbol{H}$ depends *only* on H_i, the state of nature at the epoch i when this observation is taken.

The results obtained earlier apply for it suffices to identify the state of nature as the rv $\boldsymbol{H}$ and the observation as $\boldsymbol{X}$: We then see that the ML detector for $\boldsymbol{H}$ on the basis of the observation vector $\boldsymbol{X}$ prescribes

$$\text{Choose } \widehat{\boldsymbol{H}} = (h_1, \ldots, h_n) \quad \text{iff} \quad \prod_{i=1}^{n} f_{h_i}(\boldsymbol{x}_i) \text{ largest.}$$

This leads to the following equivalent prescription

$$\text{Choose } \widehat{H}_i = h_i \quad \text{iff} \quad f_{h_i}(\boldsymbol{x}_i) \text{ largest,} \quad i = 1, \ldots, n.$$

In other words the corresponding ML detector reduces to *sequentially* applying an appropriate ML detector for deciding the state of nature H_i at epoch i on the basis of the observation $\boldsymbol{X}_i$ collected *only* at that epoch for *each* $i = 1, \ldots, n$. Of course this is a great simplification since it can be done sequentially in time.

We now turn to the MAP detector in the situation when the rvs $H_1, \ldots, H_n$ are *mutually independent* (but not necessarily identically distributed), i.e.,

$$\mathbb{P}\left[H_1 = h_1, \ldots, H_n = h_n\right] = \prod_{i=1}^{n} \mathbb{P}\left[H_i = h_i\right] \tag{1.12}$$

with $\boldsymbol{h} = (h_1, \ldots, h_n)$ in $\mathcal{L}^n$. Under this independence assumption on the prior, the MAP detector for $\boldsymbol{H}$ on the basis of the observation vector $\boldsymbol{X}$ prescribes

$$\text{Choose } \widehat{\boldsymbol{H}} = (h_1, \ldots, h_n) \quad \text{iff} \quad \prod_{i=1}^{n} \mathbb{P}\left[H_i = h_i\right] f_{h_i}(\boldsymbol{x}_i) \text{ largest.}$$

This time again, a separation occurs under the independence assumption (1.12), namely the combined prescriptions

$$\text{Choose } \widehat{H}_i = h_i \quad \text{iff} \quad \mathbb{P}\left[H_i = h_i\right] f_{h_i}(\boldsymbol{x}_i) \text{ largest,} \quad i = 1, \ldots, n.$$

Again great simplification is achieved as the MAP detector reduces to *sequentially* applying an MAP detector for deciding the state of nature H_i at epoch i on the basis of the observation $\boldsymbol{X}_i$ collected *only* at that epoch for *each* $i = 1, \ldots, n$.

1.7 Irrelevant data

When applying the ideas of Decision Theory developed in this chapter, we shall sometimes encounter the following structured situation: The observed data $\boldsymbol{X}$ admits a natural partitioning into two component vectors, say $\boldsymbol{X} = (\boldsymbol{Y}, \boldsymbol{Z})$ for rvs $\boldsymbol{Y}$ and $\boldsymbol{Z}$ which take values in $\mathbb{R}^p$ and $\mathbb{R}^q$, respectively, with $p + q = d$. To simplify the discussion, we still assume that for each $m = 1, \ldots, M$, the distribution function F_m admits a density f_m on $\mathbb{R}^d$. In that case, the distribution of the rv $\boldsymbol{Y}$ given $H = m$ also admits a density g_m given by

$$g_m(\boldsymbol{y}) = \int_{\mathbb{R}^q} f_m(\boldsymbol{y}, \boldsymbol{z})\, d\boldsymbol{z}, \quad \boldsymbol{y} \in \mathbb{R}^p.$$

It is a simple matter to check for $\boldsymbol{y}$ in $\mathbb{R}^p$ that the conditional distribution of the rv $\boldsymbol{Z}$ given $\boldsymbol{Y} = \boldsymbol{y}$ and $H = m$ admits a density, denoted $h_m(\cdot|\boldsymbol{y})$. Standard conditioning arguments readily yield

$$f_m(\boldsymbol{y}, \boldsymbol{z}) = g_m(\boldsymbol{y}) h_m(\boldsymbol{z}|\boldsymbol{y}), \quad \boldsymbol{y} \in \mathbb{R}^p,\ \boldsymbol{z} \in \mathbb{R}^q. \tag{1.13}$$

In fact, with the convention $\frac{0}{0} = 0$, we find

$$h_m(\boldsymbol{z}|\boldsymbol{y}) = \frac{f_m(\boldsymbol{y},\boldsymbol{z})}{g_m(\boldsymbol{y})}, \quad \boldsymbol{y} \in \mathbb{R}^p, \ \boldsymbol{z} \in \mathbb{R}^q. \tag{1.14}$$

Returning to the definition (1.4) of the optimal detector, we see that $\delta^\star$ prescribes

$$\widehat{H} = m \quad \text{iff} \quad p_m g_m(\boldsymbol{y}) h_m(\boldsymbol{z}|\boldsymbol{y}) \text{ largest}$$

with a tie-breaker. Therefore, if the conditional density at (1.14) were to *not* depend on m, i.e.,

$$h_1(\boldsymbol{z}|\boldsymbol{y}) = \ldots = h_M(\boldsymbol{z}|\boldsymbol{y}) =: h(\boldsymbol{z}|\boldsymbol{y}), \quad \boldsymbol{y} \in \mathbb{R}^p, \ \boldsymbol{z} \in \mathbb{R}^q \tag{1.15}$$

then (1.14) reduces to

$$\widehat{H} = m \quad \text{iff} \quad p_m g_m(\boldsymbol{y}) \text{ largest.} \tag{1.16}$$

The condition (1.15) and the resulting form (1.16) of the optimal detector suggest that knowledge of $\boldsymbol{Z}$ plays no role in developing inference of H on the basis of the pair $(\boldsymbol{Y}, \boldsymbol{Z})$, hence the terminology *irrelevant* data given to $\boldsymbol{Z}$.

In a number of cases occuring in practice, the condition (1.15) is guaranteed by the following stronger conditional independence: (i) The rvs $\boldsymbol{Y}$ and $\boldsymbol{Z}$ are mutually independent conditionally on the rv H, and (ii) the rv $\boldsymbol{Z}$ is itself independent of the rv H. In other words, for each $m = 1, \ldots, M$, it holds that

$$\begin{aligned} \mathbb{P}[\boldsymbol{Y} \le \boldsymbol{y}, \boldsymbol{Z} \le \boldsymbol{z} | H = m] &= \mathbb{P}[\boldsymbol{Y} \le \boldsymbol{y} | H = m] \, \mathbb{P}[\boldsymbol{Z} \le \boldsymbol{z} | H = m] \\ &= \mathbb{P}[\boldsymbol{Y} \le \boldsymbol{y} | H = m] \, \mathbb{P}[\boldsymbol{Z} \le \boldsymbol{z}] \end{aligned}$$

for all $\boldsymbol{y}$ and $\boldsymbol{z}$ in $\mathbb{R}^p$ and $\mathbb{R}^q$, respectively. In that case, it is plain that

$$f_m(\boldsymbol{y},\boldsymbol{z}) = g_m(\boldsymbol{y}) h(\boldsymbol{z}), \quad \boldsymbol{y} \in \mathbb{R}^p, \ \boldsymbol{z} \in \mathbb{R}^q$$

where h is the *unconditional* probability density function of $\boldsymbol{Z}$. The validity of (1.15) is now immediate with

$$h(\boldsymbol{z}|\boldsymbol{y}) = h(\boldsymbol{z}), \quad \boldsymbol{y} \in \mathbb{R}^p, \ \boldsymbol{z} \in \mathbb{R}^q.$$

1.8 Sufficient statistics

A mapping $T : \mathbb{R}^d \rightarrow \mathbb{R}^p$ is said to be a *sufficient statistic* for (estimating) H on the basis of $\boldsymbol{X}$ if the conditional distribution of $\boldsymbol{X}$ given $H = m$ and $T(\boldsymbol{X})$ does not depend on m.

The Fisher-Neyman Factorization Theorem given next provides a convenient characterization of a sufficient statistic in the framework used here.

Theorem 1.8.1 *Assume that for each $m = 1, \ldots, M$, the distribution function F_m admits a density f_m on $\mathbb{R}^d$. The mapping $T : \mathbb{R}^d \rightarrow \mathbb{R}^p$ is a sufficient statistic for estimating H on the basis of $\boldsymbol{X}$ if and only if there exist mappings $h : \mathbb{R}^d \rightarrow \mathbb{R}_+$ and $g_1, \ldots, g_M : \mathbb{R}^p \rightarrow \mathbb{R}_+$ such that*

$$f_m(\boldsymbol{x}) = h(\boldsymbol{x}) g_m(T(\boldsymbol{x})), \quad \boldsymbol{x} \in \mathbb{R}^d \tag{1.17}$$

for each $m = 1, \ldots, M$.

The usefulness of the Fisher-Neyman Factorization Theorem should be apparent: From the definition (1.4) of the optimal detector, we see that $\delta^\star$ prescribes

$$\widehat{H} = m \quad \text{iff} \quad p_m h(\boldsymbol{x}) g_m(T(\boldsymbol{x})) \text{ largest} \tag{1.18}$$

with a tie-breaker, a prescription equivalent to

$$\widehat{H} = m \quad \text{iff} \quad p_m g_m(T(\boldsymbol{x})) \text{ largest} \tag{1.19}$$

with a tie-breaker. In many applications p is much smaller than d with obvious advantages from the point of view of storage and implementation: The data $\boldsymbol{x}$ is possibly high-dimensional but after some processing, the decision concerning the state of nature can be taken on the basis of the lower-dimensional quantity $T(\boldsymbol{x})$.

The following example, already introduced in Section 1.5, should clarify the advantage of using (1.19) over (1.18): Assume the distributions $F_1, \ldots, F_M$ to be Gaussian distributions with the same invertible covariance matrix $\sigma^2 \boldsymbol{I}_d$ but with distinct means $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_M$. Further assume that

$$\boldsymbol{\mu}_m = \lambda_m \boldsymbol{\mu}, \quad m = 1, \ldots, M$$

for distinct scalars $\lambda_1, \ldots, \lambda_M$ and non-zero vector $\boldsymbol{\mu}$. Then, under these assumptions, for each $m = 1, \ldots, M$, the distribution F_m admits the density

$$f_m(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi\sigma^2)^d}} e^{-\frac{1}{2\sigma^2}\|\boldsymbol{x} - \lambda_m \boldsymbol{\mu}\|^2}, \quad \boldsymbol{x} \in \mathbb{R}^d \tag{1.20}$$

where

$$\|\boldsymbol{x} - \lambda_m \boldsymbol{\mu}\|^2 = \|\boldsymbol{x}\|^2 - 2\lambda_m \boldsymbol{x}'\boldsymbol{\mu} + \lambda_m^2 \|\boldsymbol{\mu}\|^2.$$

As a result, the density f_m can be written in the form (1.17) with

$$h(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi\sigma^2)^p}} e^{-\frac{1}{2\sigma^2}\|\boldsymbol{x}\|^2}, \quad \boldsymbol{x} \in \mathbb{R}^d$$

and

$$g_m(t) = e^{-\frac{1}{2\sigma^2}\left(-2\lambda_m t + \lambda_m^2 \|\boldsymbol{\mu}\|^2\right)}, \quad t \in \mathbb{R}.$$

It now follows from Theorem 1.8.1 that the mapping $T : \mathbb{R}^d \to \mathbb{R}$ given by

$$T(\boldsymbol{x}) := \boldsymbol{x}'\boldsymbol{\mu}, \quad \boldsymbol{x} \in \mathbb{R}^d$$

is a sufficient statistic for (estimating) H on the basis of $\boldsymbol{X}$ – Here $p = 1$ while d is arbitrary (and often very large). While the (high-dimensional) data $\boldsymbol{x}$ is observed, the decision is taken on the basis of the *one*-dimensional quantity $T(\boldsymbol{x})$, namely

$$(1.21) \qquad \widehat{H} = m \quad \text{iff} \quad \log p_m - \frac{1}{2\sigma^2}\left(-2\lambda_m T(\boldsymbol{x}) + \lambda_m^2 \|\boldsymbol{\mu}\|^2\right) \text{ largest}$$

upon taking logarithms in (1.19).

1.9 Exercises

Ex. 1.1 Consider the Bayesian hypothesis problem with an arbitrary cost function $C : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$. Revisit the arguments of Section 1.2 to identify the optimal detector.

Ex. 1.2 Show that the detector identified in Exercise 1.1 is indeed the optimal detector. Arguments similar to the ones given in Section 1.3 can be used.

Ex. 1.3 Specialize Exercise 1.2 to the case $M = 2$.

Ex. 1.4 Show that the formulations (1.7) and (1.8) are equivalent.

Ex. 1.5 In the setting of Section 1.6, show that the rv $\boldsymbol{H}$ is uniformly distributed on $\mathcal{L}^n$ if and only the rvs $H_1, \ldots, H_n$ are i.i.d. rvs, each of which is uniformly distributed on $\mathcal{L}$. Use this fact to obtain the form of the ML detector from the results derived in the second half of Section 1.6, under the assumption (1.12) on the prior.

Ex. 1.6 Consider the situation where the scalar observation X and the state of nature H are rvs related through the measurement equation

$$X = \mu_H + V$$

under the following assumptions: The rvs H and V are mutually independent, the rv H takes values in some finite set $\mathcal{H} = \{1, \ldots, M\}$, and the $\mathbb{R}$-valued rv V admits a density f_V. Here $\mu_1, \ldots, \mu_M$ denote distinct scalars, say $\mu_1 < \ldots < \mu_M$. Find the corresponding ML detector.

Ex. 1.7 Continue Exercise 1.6 when the noise V has a Cauchy distribution with density

$$f_V(v) = \frac{1}{\pi(1+v^2)}, \quad v \in \mathbb{R}.$$

Show that the ML detector implements nearest-neighbor detection.

Ex. 1.8 Consider the multi-dimensional version of Exercise 1.6 with the observation $\boldsymbol{X}$ and the state of nature H related through the measurement equation

$$\boldsymbol{X} = \boldsymbol{\mu}_H + \boldsymbol{V}$$

under the following assumptions: The rvs H and $\boldsymbol{V}$ are mutually independent, the rv H takes values in some finite set $\mathcal{H} = \{1, \ldots, M\}$, and the $\mathbb{R}^d$-valued rv $\boldsymbol{V}$ admits a density $f_{\boldsymbol{V}}$. Here the vectors $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_M$ are distinct elements of $\mathbb{R}^d$. Find the ML detector when $f_{\boldsymbol{V}}$ is of the form

$$f_{\boldsymbol{V}}(\boldsymbol{v}) = g(\|\boldsymbol{v}\|^2), \quad \boldsymbol{v} \in \mathbb{R}^d$$

for some decreasing function $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$.

Chapter 2

Gaussian Random Variables

This chapter is devoted to a brief discussion of the class of Gaussian rvs. In particular, for easy reference we have collected various facts and properties to be used repeatedly.

2.1 Scalar Gaussian rvs

With

$$\mu \in \mathbb{R} \quad \text{and} \quad \sigma \geq 0,$$

an $\mathbb{R}$-valued rv X is said to be a *Gaussian* (or normally distributed) rv with mean μ and variance σ^2 if either it is degenerate to a constant with $X = \mu$ a.s. (in which case $\sigma = 0$) or the probability distribution of X is of the form

$$\mathbb{P}\left[X \leq x\right] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt, \quad x \in \mathbb{R}$$

(in which case $\sigma^2 > 0$). Under either circumstance, it can be shown that

$$\mathbb{E}\left[e^{i\theta X}\right] = e^{i\theta\mu - \frac{\sigma^2}{2}\cdot\theta^2}, \quad \theta \in \mathbb{R}. \tag{2.1}$$

It is then follows by differentiation that

$$\mathbb{E}\left[X\right] = \mu \quad \text{and} \quad \mathbb{E}\left[X^2\right] = \mu^2 + \sigma^2 \tag{2.2}$$

so that $\mathrm{Var}[X] = \sigma^2$. This confirms the meaning ascribed to the parameters μ and σ^2 as mean and variance, respectively.

It is a simple matter to check that if X is normally distributed with mean μ and variance σ^2, then for scalars a and b, the rv $aX + b$ is also normally distributed with mean $a\mu + b$ and variance $a^2\sigma^2$. In particular, with $\sigma > 0$, the rv $\sigma^{-1}(X - \mu)$ is a Gaussian rv with mean zero and unit variance.

2.2 The standard Gaussian rv

The Gaussian rv with mean zero and unit variance occupies a very special place among Gaussian rvs, and is often referred to as the *standard* Gaussian rv. Throughout, we denote by U the Gaussian rv with zero mean and unit variance. Its probability distribution function is given by

$$\mathbb{P}\left[U \leq x\right] = \Phi(x) := \int_{-\infty}^{x} \phi(t)dt, \quad x \in \mathbb{R} \tag{2.3}$$

with density function ϕ given by

$$\phi(x) := \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}. \tag{2.4}$$

As should be clear from earlier comments, the importance of this standard rv U stems from the fact that for any Gaussian rv X with mean μ and variance σ^2, it holds that $X =_{st} \mu + \sigma U$, so that

$$\begin{aligned}
\mathbb{P}\left[X \leq x\right] &= \mathbb{P}\left[\sigma^{-1}(X - \mu) \leq \sigma^{-1}(x - \mu)\right] \\
&= \mathbb{P}\left[U \leq \sigma^{-1}(x - \mu)\right] \\
&= \Phi(\sigma^{-1}(x - \mu)), \quad x \in \mathbb{R}.
\end{aligned}$$

The evaluation of probabilities involving Gaussian rvs thus reduces to the evaluation of related probabilities for the standard Gaussian rv.

For each x in $\mathbb{R}$, we note by symmetry that $\mathbb{P}\left[U \leq -x\right] = \mathbb{P}\left[U > x\right]$, so that $\Phi(-x) = 1 - \Phi(x)$, and Φ is therefore fully determined by the complementary probability distribution function of U on $[0, \infty)$, namely

$$Q(x) := 1 - \Phi(x) = \mathbb{P}\left[U > x\right], \quad x \geq 0. \tag{2.5}$$

2.3 Gaussian integrals

There are a number of integrals that can be evaluated explicitly by making use of the fact that the Gaussian density function (2.4) must integrate to unity. We refer to these integrals as *Gaussian integrals*, and provide an expression for them.

Lemma 2.3.1 *For every a in $\mathbb{R}$ and $b > 0$, it holds that*

$$I(a,b) := \int_{\mathbb{R}} e^{ax-bx^2}\,dx = \sqrt{\frac{\pi}{b}}e^{\frac{a^2}{4b}}. \tag{2.6}$$

Proof. To evaluate $I(a,b)$ we use a "completion-of-square" argument to write

$$ax - bx^2 = -b\left(x^2 - \frac{a}{b}x\right) = -b\left(x - \frac{a}{2b}\right)^2 + \frac{a^2}{4b}, \quad x \in \mathbb{R}$$

so that

$$\begin{aligned} I(a,b) &= e^{\frac{a^2}{4b}} \int_{\mathbb{R}} e^{-b\left(x-\frac{a}{2b}\right)^2}\,dx \\ &= \sqrt{\frac{\pi}{b}}e^{\frac{a^2}{4b}} \int_{\mathbb{R}} \sqrt{\frac{b}{\pi}}e^{-b\left(x-\frac{a}{2b}\right)^2}\,dx. \end{aligned}$$

The desired conclusion (2.6) follows once we observe that

$$\int_{\mathbb{R}} \sqrt{\frac{b}{\pi}}e^{-b\left(x-\frac{a}{2b}\right)^2}\,dx = 1$$

as the integral of a Gaussian density with mean $\mu = \frac{a}{2b}$ and variance $\sigma^2 = \frac{1}{2b}$. ■

Sometimes we shall be faced with the task of evaluating integrals that reduce to integrals of the form (2.6). This is taken on in

Lemma 2.3.2 *For every pair a and b in $\mathbb{R}$, it holds that*

$$\begin{aligned} J(\lambda; a, b) &:= \int_{\mathbb{R}} e^{-\lambda(a+bx)^2}\phi(x)\,dx \\ &= \frac{1}{\sqrt{1+2\lambda b^2}} \cdot e^{-\frac{\lambda a^2}{1+2\lambda b^2}}, \quad \lambda > 0. \end{aligned} \tag{2.7}$$

Proof. Fix $\lambda > 0$. For each x in $\mathbb{R}$, we note that

$$\frac{1}{2}x^2 + \lambda(a+bx)^2 = \frac{1}{2}\left(1+2\lambda b^2\right)x^2 + \lambda a^2 + 2\lambda abx.$$

Hence, upon making the change of variable $u = x\sqrt{1+2\lambda b^2}$, we find

$$\begin{aligned} J(\lambda; a, b) &= e^{-\lambda a^2} \int_{\mathbb{R}} \phi(\sqrt{1+2\lambda b^2}x) e^{-2\lambda abx}\, dx \\ &= e^{-\lambda a^2} \int_{\mathbb{R}} e^{-\frac{2\lambda ab}{\sqrt{1+2\lambda b^2}}u} \phi(u)\, \frac{du}{\sqrt{1+2\lambda b^2}} \\ &= \frac{e^{-\lambda a^2}}{\sqrt{1+2\lambda b^2}} \int_{\mathbb{R}} e^{-\frac{2\lambda ab}{\sqrt{1+2\lambda b^2}}u} \phi(u)\, du \\ &= \frac{e^{-\lambda a^2}}{\sqrt{2\pi(1+2\lambda b^2)}} I(\alpha, \beta) \qquad (2.8) \end{aligned}$$

with

$$\alpha := -\frac{2\lambda ab}{\sqrt{1+2\lambda b^2}} \quad \text{and} \quad \beta := \frac{1}{2}.$$

Applying Lemma 2.3.1, we note that

$$\frac{\alpha^2}{4\beta} = \frac{\alpha^2}{2} = \frac{2\lambda^2 a^2 b^2}{1+2\lambda b^2}$$

so that

$$I(\alpha, \beta) = \sqrt{2\pi} e^{\frac{\alpha^2}{2}} = \sqrt{2\pi} e^{\frac{2\lambda^2 a^2 b^2}{1+2\lambda b^2}}. \qquad (2.9)$$

The desired conclusion readily follows from (2.8) and (2.9) once we observe that

$$-\lambda a^2 + \frac{2\lambda^2 a^2 b^2}{1+2\lambda b^2} = -\frac{\lambda a^2}{1+2\lambda b^2}.$$

■

As an easy corollary of Lemma 2.3.1, any Gaussian rv X with mean μ and variance σ^2 has a *moment generating function* given by

$$\mathbb{E}\left[e^{\theta X}\right] = e^{\theta\mu + \frac{\sigma^2}{2}\cdot\theta^2}, \quad \theta \in \mathbb{R}. \qquad (2.10)$$

Indeed, for each θ in $\mathbb{R}$, direct inspection shows that

$$\begin{aligned} \mathbb{E}\left[e^{\theta X}\right] &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\theta x - \frac{(x-\mu)^2}{2\sigma^2}}\, dx \\ &= e^{\theta\mu} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\theta t - \frac{t^2}{2\sigma^2}}\, dt \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{\theta\mu} I\left(\theta, \frac{1}{2\sigma^2}\right) \end{aligned}$$

where the second equality is obtained by the change of variable $t = x - \mu$, and (2.10) follows by making use of Lemma 2.3.1. Observe that (2.1) can also be obtained formally from (2.10) upon replacing θ in the latter by $i\theta$.

2.4 Evaluating $Q(x)$

The complementary distribution function (2.5) repeatedly enters the computation of various probabilities of error. Given its importance, we need to develop good approximations to $Q(x)$ over the entire range $x \geq 0$.

The error function In the literature on digital communications, probabilities of error are often expressed in terms of the so-called *error function* $\mathrm{Erf} : \mathbb{R}_+ \to \mathbb{R}$ and of its complement $\mathrm{Erfc} : \mathbb{R}_+ \to \mathbb{R}$ defined by

$$\mathrm{Erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt, \quad x \geq 0 \tag{2.11}$$

and

$$\mathrm{Erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt, \quad x \geq 0. \tag{2.12}$$

A simple change of variables ($t = \frac{u}{\sqrt{2}}$) in these integrals leads to the relationships

$$\mathrm{Erf}(x) = 2\left(\Phi(x\sqrt{2}) - \frac{1}{2}\right) \quad \text{and} \quad \mathrm{Erfc}(x) = 2Q(x\sqrt{2}),$$

so that

$$\mathrm{Erf}(x) = 1 - \mathrm{Erfc}(x), \quad x \geq 0.$$

Conversely, we also have

$$\Phi(x) = \frac{1}{2}\left(1 + \mathrm{Erf}\left(\frac{x}{\sqrt{2}}\right)\right) \quad \text{and} \quad Q(x) = \frac{1}{2}\mathrm{Erfc}\left(\frac{x}{\sqrt{2}}\right).$$

Thus, knowledge of any one of the quantities Φ, Q, Erf or Erfc is equivalent to that of the other three quantities. Although the last two quantities do not have a probabilistic interpretation, evaluating Erf is computationally more efficient. Indeed, $\mathrm{Erf}(x)$ is an integral of a positive function over the *finite* interval $[0, x]$ (and not over an infinite interval as in the other cases).

Chernoff bounds To approximate $Q(x)$ we begin with a crude bound which takes advantage of (2.10): Fix $x > 0$. For each $\theta > 0$, the usual Chernoff bound argument gives

$$\begin{aligned} \mathbb{P}\left[U > x\right] &\leq \mathbb{E}\left[e^{\theta U}\right] e^{-\theta x} \\ &= e^{-\theta x + \frac{\theta^2}{2}} \\ &= e^{-\frac{x^2}{2}} e^{\frac{(\theta - x)^2}{2}} \end{aligned} \tag{2.13}$$

where in the last equality we made use of a completion-of-square argument. The best lower bound

$$Q(x) \leq e^{-\frac{x^2}{2}}, \quad x \geq 0 \tag{2.14}$$

is achieved upon selecting $\theta = x$ in (2.13). We refer to the bound (2.14) as a Chernoff bound; it is not very accurate for small $x > 0$ since $\lim_{x \to 0} Q(x) = \frac{1}{2}$ while $\lim_{x \to 0} e^{-\frac{x^2}{2}} = 1$.

Approximating $Q(x)$ ($x \to \infty$) The Chernoff bound shows that $Q(x)$ decays to zero for large x at least as fast as $e^{-\frac{x^2}{2}}$. However, sometimes more precise information is needed regarding the rate of decay of $Q(x)$. This issue is addressed as follows:

For each $x \geq 0$, a straigthforward change of variable yields

$$\begin{aligned} Q(x) &= \int_x^\infty \phi(t) dt \\ &= \int_0^\infty \phi(x + t) dt \\ &= \phi(x) \int_0^\infty e^{-xt} e^{-\frac{t^2}{2}} dt. \end{aligned} \tag{2.15}$$

With the Taylor series expansion of $e^{-\frac{t^2}{2}}$ in mind, approximations for $Q(x)$ of increased accuracy thus suggest themselves by simply approximating the second exponential factor (namely e^{-xt}) in the integral at (2.15) by terms of the form

$$\sum_{k=0}^{n} \frac{(-1)^k}{2^k k!} t^{2k}, \quad n = 0, 1, \ldots \tag{2.16}$$

To formulate the resulting approximation contained in Proposition 2.4.1 given next, we set

$$Q_n(x) = \phi(x) \int_0^\infty \left(\sum_{k=0}^n \frac{(-1)^k}{2^k k!} t^{2k} \right) e^{-xt} dt, \quad x \geq 0$$

for each $n = 0, 1, \ldots$.

Proposition 2.4.1 *Fix* $n = 0, 1, \ldots$. *For each* $x > 0$ *it holds that*

$$Q_{2n+1}(x) \leq Q(x) \leq Q_{2n}(x), \tag{2.17}$$

with

$$\mid Q(x) - Q_n(x) \mid \leq \frac{(2n)!}{2^n n!} x^{-(2n+1)} \phi(x). \tag{2.18}$$

where

$$Q_n(x) = \phi(x) \sum_{k=0}^n \frac{(-1)^k (2k)!}{2^k k!} x^{-(2k+1)}. \tag{2.19}$$

A proof of Proposition 2.4.1 can be found in Section 2.12. Upon specializing (2.17) to $n = 0$ we get

$$\frac{e^{-\frac{x^2}{2}}}{x\sqrt{2\pi}} \left(1 - \frac{1}{x^2} \right) \leq Q(x) \leq \frac{e^{-\frac{x^2}{2}}}{x\sqrt{2\pi}}, \quad x > 0 \tag{2.20}$$

and the asymptotics

$$Q(x) \sim \frac{e^{-\frac{x^2}{2}}}{x\sqrt{2\pi}} \quad (x \to \infty) \tag{2.21}$$

follow. Note that the lower bound in (2.20) is meaningful only when $x \geq 1$.

2.5 Gaussian random vectors

Let $\boldsymbol{\mu}$ denote a vector in $\mathbb{R}^d$ and let $\boldsymbol{\Sigma}$ be a symmetric and non-negative definite $d \times d$ matrix, i.e., $\boldsymbol{\Sigma}' = \boldsymbol{\Sigma}$ and $\boldsymbol{\theta}'\boldsymbol{\Sigma}\boldsymbol{\theta} \geq 0$ for all $\boldsymbol{\theta}$ in $\mathbb{R}^d$.

An $\mathbb{R}^d$-valued rv $\boldsymbol{X}$ is said to be a Gaussian rv with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ if there exist a $d \times p$ matrix $\boldsymbol{T}$ for some positive integer p and i.i.d. zero mean unit variance Gaussian rvs $U_1, \ldots, U_p$ such that

$$\boldsymbol{T}\boldsymbol{T}' = \boldsymbol{\Sigma} \tag{2.22}$$

and

$$\boldsymbol{X} =_{st} \boldsymbol{\mu} + \boldsymbol{T}\boldsymbol{U}_p \tag{2.23}$$

where $\boldsymbol{U}_p$ is the $\mathbb{R}^p$-valued rv $(U_1, \ldots, U_p)'$.

From (2.22) and (2.23) it is plain that

$$\mathbb{E}\left[\boldsymbol{X}\right] = \mathbb{E}\left[\boldsymbol{\mu} + \boldsymbol{T}\boldsymbol{U}_p\right] = \boldsymbol{\mu} + \boldsymbol{T}\mathbb{E}\left[\boldsymbol{U}_p\right] = \boldsymbol{\mu}$$

and

$$\begin{aligned} \mathbb{E}\left[(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})'\right] &= \mathbb{E}\left[\boldsymbol{T}\boldsymbol{U}_p(\boldsymbol{T}\boldsymbol{U}_p)'\right] \\ &= \boldsymbol{T}\mathbb{E}\left[\boldsymbol{U}_p\boldsymbol{U}_p'\right]\boldsymbol{T}' \\ &= \boldsymbol{T}\boldsymbol{I}_p\boldsymbol{T}' = \boldsymbol{\Sigma}, \end{aligned} \tag{2.24}$$

whence

$$\mathbb{E}\left[\boldsymbol{X}\right] = \boldsymbol{\mu} \quad \text{and} \quad \mathrm{Cov}[\boldsymbol{X}] = \boldsymbol{\Sigma}.$$

Again this confirms the terminology used for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ as mean vector and covariance matrix, respectively.

It is a well-known fact from Linear Algebra [, , p.] that for any symmetric and non-negative definite $d \times d$ matrix $\boldsymbol{\Sigma}$, there exists a $d \times d$ matrix $\boldsymbol{T}$ such that (2.22) holds with $p = d$. This matrix $\boldsymbol{T}$ can be selected to be *symmetric* and *non-negative definite*, and is called the *square root* of $\boldsymbol{\Sigma}$. Consequently, for any vector $\boldsymbol{\mu}$ in $\mathbb{R}^d$ and any symmetric non-negative definite $d \times d$ matrix $\boldsymbol{\Sigma}$, there always exists an $\mathbb{R}^d$-valued Gaussian rv $\boldsymbol{X}$ with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ – Simply take

$$\boldsymbol{X} =_{st} \boldsymbol{\mu} + \boldsymbol{T}\boldsymbol{U}_d$$

where $\boldsymbol{T}$ is the square root of $\boldsymbol{\Sigma}$.

2.6 Characteristic functions

The characteristic function of Gaussian rvs has an especially simple form which is now developed.

Lemma 2.6.1 *The characteristic function of a Gaussian $\mathbb{R}^d$-valued rv $\boldsymbol{X}$ with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ is given by*

$$\mathbb{E}\left[e^{i\boldsymbol{\theta}'\boldsymbol{X}}\right] = e^{i\boldsymbol{\theta}'\boldsymbol{\mu} - \frac{1}{2}\boldsymbol{\theta}'\boldsymbol{\Sigma}\boldsymbol{\theta}}, \quad \boldsymbol{\theta} \in \mathbb{R}^d. \tag{2.25}$$

Conversely, any $\mathbb{R}^d$-valued rv $\boldsymbol{X}$ whose characteristic function is given by (2.25) for some vector $\boldsymbol{\mu}$ in $\mathbb{R}^d$ and symmetric non-negative definite $d \times d$ matrix $\boldsymbol{\Sigma}$ is a Gaussian $\mathbb{R}^d$-valued rv $\boldsymbol{X}$ with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.

Proof. Consider an $\mathbb{R}^d$-valued rv $\boldsymbol{X}$ which is a Gaussian rv with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. By definition, there exist a $d \times p$ matrix $\boldsymbol{T}$ for some positive integer p and i.i.d. zero mean unit variance Gaussian rvs $U_1, \ldots, U_p$ such that (2.22) and (2.23) hold.

For each $\boldsymbol{\theta}$ in $\mathbb{R}^d$, we get

$$\begin{aligned}
\mathbb{E}\left[e^{i\boldsymbol{\theta}'\boldsymbol{X}}\right] &= e^{i\boldsymbol{\theta}'\boldsymbol{\mu}} \cdot \mathbb{E}\left[e^{i\boldsymbol{\theta}'\boldsymbol{T}\boldsymbol{U}_p}\right] \\
&= e^{i\boldsymbol{\theta}'\boldsymbol{\mu}} \cdot \mathbb{E}\left[e^{i(\boldsymbol{T}'\boldsymbol{\theta})'\boldsymbol{U}_p}\right] \\
&= e^{i\boldsymbol{\theta}'\boldsymbol{\mu}} \cdot \mathbb{E}\left[e^{i\sum_{k=1}^p (\boldsymbol{T}'\boldsymbol{\theta})_k U_k}\right] \\
&= e^{i\boldsymbol{\theta}'\boldsymbol{\mu}} \cdot \prod_{k=1}^p \mathbb{E}\left[e^{i(\boldsymbol{T}'\boldsymbol{\theta})_k U_k}\right] \qquad (2.26) \\
&= e^{i\boldsymbol{\theta}'\boldsymbol{\mu}} \cdot \prod_{k=1}^p e^{-\frac{1}{2}|(\boldsymbol{T}'\boldsymbol{\theta})_k|^2} \qquad (2.27)
\end{aligned}$$

The equality (2.26) is a consequence of the independence of the rvs $U_1, \ldots, U_p$, while (2.27) follows from their Gaussian character (and (2.1)).

Next, we note that

$$\begin{aligned}
\sum_{k=1}^p |(\boldsymbol{T}'\boldsymbol{\theta})_k|^2 &= (\boldsymbol{T}'\boldsymbol{\theta})'(\boldsymbol{T}'\boldsymbol{\theta}) \\
&= \boldsymbol{\theta}'(\boldsymbol{T}\boldsymbol{T}')\boldsymbol{\theta} = \boldsymbol{\theta}'\boldsymbol{\Sigma}\boldsymbol{\theta} \qquad (2.28)
\end{aligned}$$

upon invoking (2.22). It is now plain from (2.27) that the characteristic function of the Gaussian $\mathbb{R}^d$-valued rv $\boldsymbol{X}$ is given by (2.25).

Conversely, consider an $\mathbb{R}^d$-valued rv $\boldsymbol{X}$ with characteristic function of the form (2.25) for some vector $\boldsymbol{\mu}$ in $\mathbb{R}^d$ and some symmetric non-negative definite $d \times d$ matrix $\boldsymbol{\Sigma}$. By comments made earlier, there exists a $d \times d$ matrix $\boldsymbol{T}$ such that (2.22) holds. By the first part of the proof, the $\mathbb{R}^d$-valued rv $\widetilde{\boldsymbol{X}}$ given by $\widetilde{\boldsymbol{X}} := \boldsymbol{\mu} + \boldsymbol{T}\boldsymbol{U}_d$ has characteristic function given by (2.25). Since a probability distribution is completely determined by its characteristic function, it follows that the rvs $\boldsymbol{X}$ and $\widetilde{\boldsymbol{X}}$ obey the same distribution. The rv $\widetilde{\boldsymbol{X}}$ being Gaussian with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, the rv $\boldsymbol{X}$ is necessarily Gaussian as well with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. ■

2.7 Existence of a density

In general, an $\mathbb{R}^d$-valued Gaussian rv as defined above may not admit a density function. To see why, consider the null space of its covariance matrix $\boldsymbol{\Sigma}$,[1] namely

$$N(\boldsymbol{\Sigma}) := \{\boldsymbol{x} \in \mathbb{R}^d : \ \boldsymbol{\Sigma x} = \mathbf{0}_d\}.$$

Observe that $\boldsymbol{\theta}'\boldsymbol{\Sigma\theta} = 0$ if and only if $\boldsymbol{\theta}$ belongs to $N(\boldsymbol{\Sigma})$, in which case (2.25) yields

$$\mathbb{E}\left[e^{i\boldsymbol{\theta}'(\boldsymbol{X}-\boldsymbol{\mu})}\right] = 1$$

and we conclude that

$$\boldsymbol{\theta}'(\boldsymbol{X}-\boldsymbol{\mu}) = 0 \quad \text{a.s.}$$

In other words, with probability one, the rv $\boldsymbol{X}-\boldsymbol{\mu}$ is orthogonal to the linear space $N(\boldsymbol{\Sigma})$.

To proceed, we assume that the covariance matrix $\boldsymbol{\Sigma}$ is not trivial (in that it has some non-zero entries) for otherwise $\boldsymbol{X} = \boldsymbol{\mu}$ a.s. In the non-trivial case, there are now two possibilities depending on the $d \times d$ matrix $\boldsymbol{\Sigma}$ being positive definite or not. Note that the positive definiteness of $\boldsymbol{\Sigma}$, i.e., $\boldsymbol{\theta}'\boldsymbol{\Sigma\theta} = 0$ necessarily implies $\boldsymbol{\theta} = \mathbf{0}_d$, is equivalent to the condition $N(\boldsymbol{\Sigma}) = \{\mathbf{0}_d\}$.

If the $d \times d$ matrix $\boldsymbol{\Sigma}$ is not positive definite, hence only positive semi-definite, then the mass of the rv $\boldsymbol{X}-\boldsymbol{\mu}$ is concentrated on the orthogonal space $N(\boldsymbol{\Sigma})^\perp$ of $N(\boldsymbol{\Sigma})$, whence the distribution of $\boldsymbol{X}$ has its support on the linear manifold $\boldsymbol{\mu} + N(\boldsymbol{\Sigma})^\perp$ and is singular with respect to Lebesgue measure.

On the other hand, if the $d \times d$ matrix $\boldsymbol{\Sigma}$ is positive definite, then the matrix $\boldsymbol{\Sigma}$ is invertible, $\det(\boldsymbol{\Sigma}) \neq 0$ and the Gaussian rv $\boldsymbol{X}$ with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ admits a density function given by

$$f(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^d \det(\boldsymbol{\Sigma})}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}, \quad \boldsymbol{x} \in \mathbb{R}^d.$$

[1] This linear space is sometimes called the kernel of $\boldsymbol{\Sigma}$.

2.8 Linear transformations

The following result is very useful in many contexts, and shows that linear transformations preserve the Gaussian character:

Lemma 2.8.1 *let $\boldsymbol{\nu}$ be an element of $\mathbb{R}^q$ and let $\boldsymbol{A}$ be an $q \times d$ matrix. Then, for any Gaussian rv $\mathbb{R}^d$-valued rv $\boldsymbol{X}$ with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, the $\mathbb{R}^q$-valued rv $\boldsymbol{Y}$ given by*

$$\boldsymbol{Y} = \boldsymbol{\nu} + \boldsymbol{A}\boldsymbol{X}$$

is also a Gaussian rv with mean vector $\boldsymbol{\nu} + \boldsymbol{A}\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{A}\Sigma\boldsymbol{A}'$.

Proof. First, by linearity we note that

$$\mathbb{E}\left[\boldsymbol{Y}\right] = \mathbb{E}\left[\boldsymbol{\nu} + \boldsymbol{A}\boldsymbol{X}\right] = \boldsymbol{\nu} + \boldsymbol{A}\boldsymbol{\mu}$$

so that

$$\begin{aligned} \mathrm{Cov}[\boldsymbol{Y}] &= \mathbb{E}\left[\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{\mu})\left(\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{\mu})\right)'\right] \\ &= \boldsymbol{A}\mathbb{E}\left[(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})'\right]\boldsymbol{A}' \\ &= \boldsymbol{A}\Sigma\boldsymbol{A}'. \end{aligned} \tag{2.29}$$

Consequently, the $\mathbb{R}^q$-valued rv $\boldsymbol{Y}$ has mean vector $\boldsymbol{\nu}+\boldsymbol{A}\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{A}\Sigma\boldsymbol{A}'$.

Next, by the Gaussian character of $\boldsymbol{X}$, there exist a $d \times p$ matrix $\boldsymbol{T}$ for some positive integer p and i.i.d. zero mean unit variance Gaussian rvs $U_1, \ldots, U_p$ such that (2.22) and (2.23) hold. Thus,

$$\begin{aligned} \boldsymbol{Y} &=_{st} \boldsymbol{\nu} + \boldsymbol{A}\left(\boldsymbol{\mu} + \boldsymbol{T}\boldsymbol{U}_p\right) \\ &= \boldsymbol{\nu} + \boldsymbol{A}\boldsymbol{\mu} + \boldsymbol{A}\boldsymbol{T}\boldsymbol{U}_p \\ &= \widetilde{\boldsymbol{\mu}} + \widetilde{\boldsymbol{T}}\boldsymbol{U}_p \end{aligned} \tag{2.30}$$

with

$$\widetilde{\boldsymbol{\mu}} := \boldsymbol{\nu} + \boldsymbol{A}\boldsymbol{\mu} \quad \text{and} \quad \widetilde{\boldsymbol{T}} := \boldsymbol{A}\boldsymbol{T}$$

and the Gaussian character of $\boldsymbol{Y}$ is established. ■

This result can also be established through the evaluation of the characteristic function of the rv $\boldsymbol{Y}$. As an immediate consequence of Lemma 2.8.1 we get

Corollary 2.8.1 *Consider a Gaussian rv $\mathbb{R}^d$-valued rv $\boldsymbol{X}$ with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. For any subset I of $\{1, \ldots, d\}$ with $|I| = q \leq d$, the $\mathbb{R}^q$-valued rv $\boldsymbol{X}_I$ given by $\boldsymbol{X}_I = (X_i,\ i \in I)'$ is a Gaussian rv with mean vector $(\mu_i,\ i \in I)'$ and covariance matrix $(\Sigma_{ij},\ i, j \in I)$.*

2.9 Independence of Gaussian rvs

Characterizing the mutual independence of Gaussian rvs turns out to be quite straightforward as the following suggests: Consider the rvs $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_r$ where for each $s = 1, \ldots, r$, the rv$\boldsymbol{X}_s$ is an $\mathbb{R}^{d_s}$-valued rv with mean vector $\boldsymbol{\mu}_s$ and covariance matrix $\boldsymbol{\Sigma}_s$. With $d = d_1 + \ldots + d_r$, let $\boldsymbol{X}$ denote the $\mathbb{R}^d$-valued rv obtained by concatenating $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_r$, namely

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \vdots \\ \boldsymbol{X}_r \end{pmatrix}. \tag{2.31}$$

Its mean vector $\boldsymbol{\mu}$ is simply

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \vdots \\ \boldsymbol{\mu}_r \end{pmatrix} \tag{2.32}$$

while its covariance matrix $\boldsymbol{\Sigma}$ can be written in block form as

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_{1,2} & \ldots & \boldsymbol{\Sigma}_{1,r} \\ \boldsymbol{\Sigma}_{2,1} & \boldsymbol{\Sigma}_2 & \ldots & \boldsymbol{\Sigma}_{2,r} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{\Sigma}_{r,1} & \boldsymbol{\Sigma}_{r,2} & \ldots & \boldsymbol{\Sigma}_r \end{pmatrix} \tag{2.33}$$

with the notation

$$\boldsymbol{\Sigma}_{s,t} := \mathrm{Cov}[\boldsymbol{X}_s, \boldsymbol{X}_t] \quad s, t = 1, \ldots, r.$$

Lemma 2.9.1 *With the notation above, assume the $\mathbb{R}^d$-valued rv $\boldsymbol{X}$ to be a Gaussian rv with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Then, for each $s = 1, \ldots, r$, the rv $\boldsymbol{X}_s$ is a Gaussian rv with mean vector $\boldsymbol{\mu}_s$ and covariance matrix $\boldsymbol{\Sigma}_s$. Moreover, the rvs $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_r$ are mutually independent Gaussian rvs if and only they are uncorrelated, i.e.,*

$$\boldsymbol{\Sigma}_{s,t} = \delta(s,t)\boldsymbol{\Sigma}_t, \quad s, t = 1, \ldots, r. \tag{2.34}$$

The first part of Lemma 2.9.1 is a simple rewrite of Corollary 2.8.1. Sometimes we refer to the fact that the rv $\boldsymbol{X}$ is Gaussian by saying that the rvs $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_r$ are *jointly* Gaussian. A converse to Lemma 2.9.1 is available:

Lemma 2.9.2 *Assume that for each $s = 1, \ldots, r$, the rv $\boldsymbol{X}_s$ is a Gaussian rv with mean vector $\boldsymbol{\mu}_s$ and covariance matrix Σ_s. If the rvs $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_r$ are mutually independent, then the $\mathbb{R}^d$-valued rv $\boldsymbol{X}$ is an $\mathbb{R}^d$-valued Gaussian rv with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ as given by (2.33) with (2.34).*

It might be tempting to conclude that the Gaussian character of *each* of the rvs $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_r$ *alone* suffices to imply the Gaussian character of the combined rv $\boldsymbol{X}$. However, it can be shown through simple counterexamples that this is not so. In other words, the joint Gaussian character of $\boldsymbol{X}$ does not follow merely from that of its components $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_r$ *without* further assumptions.

2.10 Convergence and limits of Gaussian rvs

In later chapters we will need to define integrals with respect to Gaussian processes. As in the deterministic case, these *stochastic* integrals will be defined as limits of partial sums of the form

$$X_n := \sum_{i=1}^{k_n} a_j^{(n)} Y_j^{(n)}, \quad n = 1, 2, \ldots \tag{2.35}$$

where for each $n = 1, 2, \ldots$, the integer k_n and the coefficients $a_j^{(n)}$, $j = 1, \ldots, k_n$, are non-random while the rvs $\{Y_j^{(n)},\ j = 1, \ldots, k_n\}$ are *jointly* Gaussian rvs. Typically, as n goes to infinity so does k_n. Note that under the foregoing assumptions for each $n = 1, 2, \ldots$, the rv X_n is Gaussian with

$$\mathbb{E}\left[X_n\right] = \sum_{i=1}^{k_n} a_j^{(n)} \mathbb{E}\left[Y_j^{(n)}\right] \tag{2.36}$$

and

$$\mathrm{Var}[X_n] = \sum_{i=1}^{k_n} \sum_{j=1}^{k_n} a_i^{(n)} a_j^{(n)} \mathrm{Cov}[Y_i^{(n)}, Y_j^{(n)}]. \tag{2.37}$$

Therefore, the study of such integrals is expected to pass through the convergence of sequence of rvs $\{X_n,\ n = 1, 2, \ldots\}$ of the form (2.35). Such considerations lead naturally to the need for the following result [, Thm. , p.]:

Lemma 2.10.1 *Let $\{\boldsymbol{X}_k,\ k = 1, 2, \ldots\}$ denote a collection of $\mathbb{R}^d$-valued Gaussian rvs. For each $k = 1, 2, \ldots$, let $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ denotes the mean vector and covariance matrix of the rv $\boldsymbol{X}_k$. The rvs $\{\boldsymbol{X}_k,\ k = 1, \ldots\}$ converge in distribution (in law) if and only there exist an element $\boldsymbol{\mu}$ in $\mathbb{R}^d$ and a $d \times d$ matrix $\boldsymbol{\Sigma}$ such that*

$$\lim_{k\to\infty} \boldsymbol{\mu}_k = \boldsymbol{\mu} \quad \text{and} \quad \lim_{k\to\infty} \boldsymbol{\Sigma}_k = \boldsymbol{\Sigma}. \tag{2.38}$$

In that case,

$$\boldsymbol{X}_k \Longrightarrow_k \boldsymbol{X}$$

where $\boldsymbol{X}$ is an $\mathbb{R}^d$-valued Gaussian rv with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.

The second half of condition (2.38) ensures that the matrix $\boldsymbol{\Sigma}$ is symmetric and non-negative definite, hence a covariance matrix.

Returning to the partial sums (2.35) we see that Lemma 2.10.1 (applied with $d = 1$) requires identifying the limits $\mu = \lim_{n\to\infty} \mathbb{E}\left[X_n\right]$ and $\sigma^2 = \lim_{n\to\infty} \mathrm{Var}[X_n]$, in which case $X_n \Longrightarrow_n X$ where X is an $\mathbb{R}$-valued Gaussian rv with mean μ and variance Σ. In Section **??** we discuss a situation where this can be done quite easily.

2.11 Rvs derived from Gaussian rvs

Rayleigh rvs A rv X is said to be a *Rayleigh* rv with parameter σ ($\sigma > 0$) if

$$X =_{st} \sqrt{Y^2 + Z^2} \tag{2.39}$$

with Y and Z independent zero mean Gaussian rvs with variance σ^2. It is easy to check that

$$\mathbb{P}\left[X > x\right] = e^{-\frac{x^2}{2\sigma^2}}, \quad x \geq 0 \tag{2.40}$$

with corresponding density function

$$\frac{d}{dx}\mathbb{P}\left[X \leq x\right] = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, \quad x \geq 0. \tag{2.41}$$

It is also well known that the rv Θ given by

$$\Theta := \arctan\left(\frac{Z}{Y}\right) \tag{2.42}$$

is uniformly distributed over $[0, 2\pi)$ and independent of the Rayleigh rv X, i.e.,

$$\mathbb{P}\left[X \le x, \Theta \le \theta\right] = \frac{\theta}{2\pi}\left(1 - e^{-\frac{x^2}{2\sigma^2}}\right), \quad \theta \in [0, 2\pi),\ x \ge 0. \tag{2.43}$$

Rice rvs A rv X is said to be a *Rice* rv with parameters α (in $\mathbb{R}$) and σ ($\sigma > 0$) if

$$X =_{st} \sqrt{(\alpha + Y)^2 + Z^2} \tag{2.44}$$

with Y and Z independent zero mean Gaussian rvs with variance σ^2. It is easy to check that X admits a probability density function given by

$$\frac{d}{dx}\mathbb{P}\left[X \le x\right] = \frac{x}{\sigma^2}e^{-\frac{x^2+\alpha^2}{2\sigma^2}} \cdot I_0\left(\frac{\alpha x}{\sigma^2}\right), \quad x \ge 0. \tag{2.45}$$

Here,

$$I_0(x) := \frac{1}{2\pi}\int_0^{2\pi} e^{x\cos t}dt, \quad x \in \mathbb{R} \tag{2.46}$$

is the modified Bessel function of the first kind of order zero.

Chi-square rvs For each $n = 1, 2, \ldots$, the Chi-square rv with n degrees of freedom is the rv defined by

$$\chi_n^2 =_{st} U_1^2 + \ldots + U_n^2$$

where $U_1, \ldots, U_n$ are n i.i.d. standard Gaussian rvs.

2.12 A Proof of Proposition 2.4.1

The main idea is to use the Taylor series approximations (2.16) in the relation (2.15). To do so, we begin by establishing some elementary facts concerning the Taylor series approximations of the negative exponential e^{-y} ($y \ge 0$): For each $n = 0, 1, \ldots$, set

$$H_n(y) := \sum_{k=0}^{n} \frac{(-1)^k}{k!}y^k, \quad y \ge 0. \tag{2.47}$$

Lemma 2.12.1 *For each $y \geq 0$ and $n = 0, 1, \ldots$, it holds that*

$$H_{2n+1}(y) \leq e^{-y} \leq H_{2n}(y) \tag{2.48}$$

with

$$\mid H_n(y) - e^{-y} \mid \leq \frac{y^n}{n!}. \tag{2.49}$$

Proof. Fix $y \geq 0$ and $n = 0, 1, \ldots$. By differentiation we readily check that

$$H'_{n+1}(y) = -H_n(y),$$

so that

$$\frac{d}{dy}\left(e^{-y} - H_{n+1}(y)\right) = -\left(e^{-y} - H_n(y)\right).$$

Integrating and using the fact $H_{n+1}(0) = 1$, we find

$$e^{-y} - H_{n+1}(y) = -\int_0^y \left(e^{-t} - H_n(t)\right) dt. \tag{2.50}$$

An easy induction argument now yields (2.48) once we note for the basis step that $H_0(y) > e^{-y}$ for all $y > 0$.

To obtain the bound (2.49) on the accuracy of approximating e^{-y} by $H_n(y)$, we proceed by induction on n. For $n = 0$, it is always the case that $|e^{-y} - H_0(y)| \leq 1$, whence (2.49) holds for all $y \geq 0$ and the basis step is established. Next, we assume that (2.49) holds for all $y \geq 0$ for $n = m$ with some $m = 0, 1, \ldots$, namely

$$|e^{-y} - H_m(y)| \leq \frac{y^m}{m!}, \quad y \geq 0. \tag{2.51}$$

Hence, upon invoking (2.50) we observe that

$$\begin{aligned} |e^{-y} - H_{m+1}(y)| &\leq \int_0^y |e^{-t} - H_m(t)| dt \\ &\leq \int_0^y \frac{t^m}{m!} dt = \frac{y^{m+1}}{(m+1)!}, \quad y \geq 0 \end{aligned}$$

and the induction step is established. ■

Back to the proof of Proposition 2.4.1: Fix $x > 0$ and $n = 0, 1, \ldots$. As we have in mind to use (2.48) to bound the second exponential factor in the integrand of (2.15), we note that

$$\begin{aligned}\int_0^\infty e^{-xt} H_n\left(\frac{t^2}{2}\right) dt &= \sum_{k=0}^{n} \frac{(-1)^k}{2^k k!} \int_0^\infty t^{2k} e^{-xt} dt \\ &= \sum_{k=0}^{n} \frac{(-1)^k}{2^k k!} x^{-(2k+1)} \int_0^\infty u^{2k} e^{-u} du \\ &= \sum_{k=0}^{n} \frac{(-1)^k (2k)!}{2^k k!} x^{-(2k+1)} \qquad (2.52)\end{aligned}$$

where the last equality made use of the well-known closed-form expressions

$$\int_0^\infty u^p e^{-u} du = p!, \quad p = 0, 1, \ldots$$

for the moments of a standard exponential distribution.

The bounds (2.48) together with (2.15) yield the inequalities

$$\phi(x) \int_0^\infty e^{-xt} H_{2n+1}\left(\frac{t^2}{2}\right) dt \leq Q(x) \leq \phi(x) \int_0^\infty e^{-xt} H_{2n}\left(\frac{t^2}{2}\right) dt,$$

and (2.17) follows from the evaluation (2.52).

Using the definition of $Q(x)$ and $Q_n(x)$ we conclude from (2.49) that

$$\begin{aligned}\mid Q(x) - Q_n(x) \mid &= \phi(x) \left| \int_0^\infty e^{-xt} \left[e^{-\frac{t^2}{2}} - H_n\left(\frac{t^2}{2}\right)\right] dt \right| \\ &\leq \phi(x) \int_0^\infty e^{-xt} \frac{t^{2n}}{2^n n!} dt,\end{aligned}$$

and (2.18) follows. ■

2.13 Exercises

Ex. 2.1 Derive the relationships between the quantities Φ, Q, Erf or Erfc which are given in Section 2.4.

Ex. 2.2 Given the covariance matrix $\boldsymbol{\Sigma}$, explain why the representation (2.22)–(2.23) may not be unique. Give a counterexample.

Ex. 2.3 Give a proof for Lemma 2.9.1 and of Lemma 2.9.2.

Ex. 2.4 Construct an $\mathbb{R}^2$-valued rv $\boldsymbol{X} = (X_1, X_2)$ such that the $\mathbb{R}$-valued rvs X_1 and X_2 are each Gaussian but the $\mathbb{R}^2$-valued rv $\boldsymbol{X}$ is not (jointly) Gaussian.

Ex. 2.5 Derive the probability distribution function (2.40) of a Rayleigh rv with parameter σ ($\sigma > 0$).

Ex. 2.6 Show by direct arguments that if X is a Rayleigh distribution with parameter σ, then X^2 is exponentially distributed with parameter $(2\sigma^2)^{-1}$ [Hint: Compute $\mathbb{E}\left[e^{-\theta X^2}\right]$ for a Rayleigh rv X for $\theta \geq 0$.]

Ex. 2.7 Derive the probability distribution function (2.45) of a Rice rv with parameters α (in $\mathbb{R}$) and σ ($\sigma > 0$).

Ex. 2.8 Write a program to evaluate $Q_n(x)$.

Ex. 2.9 Let $X_1, \ldots, X_n$ be i.i.d. Gaussian rvs with zero mean and unit variance and write $S_n = X_1 + \ldots + X_n$. For each $a > 0$ show that

$$\mathbb{P}\left[S_n > na\right] \sim \frac{e^{-\frac{na^2}{2}}}{a\sqrt{2\pi n}} \quad (n \to \infty). \tag{2.53}$$

This asymptotic is known as the Bahadur-Rao correction to the large deviations asymptotics of S_n.

Ex. 2.10 Find all the moments $\mathbb{E}\left[U^p\right]$ ($p = 1, \ldots$) where U is a zero-mean unit variance Gaussian rv.

Ex. 2.11 Find all the moments $\mathbb{E}\left[U^p\right]$ ($p = 1, \ldots$) where X is a χ^2_n-rv with n degrees of freedom.

Chapter 3

Vector space methods

In this chapter we develop elements of the theory of vector spaces. As we shall see in subsequent chapters, vector space methods will prove useful in handling the so-called waveform channels by transforming them into vector channels. Vector spaces provide a unifying abstraction to carry out this translation. Additional information can be found in the references [**?**, **?**].

3.1 Vector spaces – Definitions

We begin by introducing the notion of *vector space*. Consider a set V whose elements are called *vectors* while we refer to the elements of $\mathbb{R}$ as *scalars*. We assume that V is equipped with an internal operation of *addition*, say $+ : V \times V \to V$, with the property that $(V, +)$ is a *commutative* group. This means that

1. (Commutativity)
$$\boldsymbol{v} + \boldsymbol{w} = \boldsymbol{w} + \boldsymbol{v}, \quad \boldsymbol{v}, \boldsymbol{w} \in V$$

2. (Associativity)
$$(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}), \quad \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$$

3. (Existence of a zero vector) There exists an element $\mathbf{0}$ in V such that
$$\boldsymbol{v} + \mathbf{0} = \boldsymbol{v} = \mathbf{0} + \boldsymbol{v}, \quad \boldsymbol{v} \in V$$

4. (Existence of negative vectors) For every vector $\boldsymbol{v}$ in V, there exists a vector in V, denoted $-\boldsymbol{v}$, such that

$$\boldsymbol{v} + (-\boldsymbol{v}) = \mathbf{0} = (-\boldsymbol{v}) + \boldsymbol{v}$$

It is a simple matter to check that there can be only one such zero vector $\mathbf{0}$, and that for every vector $\boldsymbol{v}$ in V, its negative $-\boldsymbol{v}$ is unique.

In order for the group $(V, +)$ to become a *vector space on* $\mathbb{R}$ we need to endow it with an external multiplication operation whereby multiplying a vector by a scalar is given a meaning as a vector. This *multiplication* operation, say $\cdot : \mathbb{R} \times V \to V$, is required to satisfy the following properties:

1. (Distributivity)

$$(a + b) \cdot \boldsymbol{v} = a \cdot \boldsymbol{v} + b \cdot \boldsymbol{v}, \quad a, b \in \mathbb{R},\ \boldsymbol{v} \in V$$

2. (Distributivity)

$$a \cdot (\boldsymbol{v} + \boldsymbol{w}) = a \cdot \boldsymbol{v} + a \cdot \boldsymbol{w}, \quad a \in \mathbb{R},\ \boldsymbol{v}, \boldsymbol{w} \in V$$

3. (Associativity)

$$a \cdot (b \cdot \boldsymbol{v}) = (ab) \cdot \boldsymbol{v} = b \cdot (a \cdot \boldsymbol{v}), \quad a, b \in \mathbb{R},\ \boldsymbol{v} \in V$$

4. (Unity law)

$$1 \cdot \boldsymbol{v} = \boldsymbol{v}, \quad \boldsymbol{v} \in V$$

It is customary to drop the multiplication symbol $\cdot$ from the notation, as we do from now. Two important examples will be developed in Chapter 4, namely the usual space $\mathbb{R}^d$ and the space of finite energy signals defined on some interval.

Throughout the remainder of this chapter, we assume given a vector space $(V, +)$ on $\mathbb{R}$.

3.2 Linear independence

Given a *finite* collection of vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ in V, the vector $\sum_{i=1}^p a_i \boldsymbol{v}_i$ is called a *linear combination* of the vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ in V (with weights $a_1, \ldots, a_p$ in $\mathbb{R}$).

The vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ in V are *linearly independent* if the relation

$$\sum_{i=1}^{p} a_i \boldsymbol{v}_i = \mathbf{0} \tag{3.1}$$

with scalars $a_1, \ldots, a_p$ in $\mathbb{R}$ implies

$$a_1 = \ldots = a_p = 0. \tag{3.2}$$

In that case, we necessarily have $\boldsymbol{v}_i \neq \mathbf{0}$ for each $i = 1, 2, \ldots, p$ (for otherwise (3.1) does not necessarily imply (3.2)).

If the vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ are linearly independent in V, then the relation

$$\sum_{i=1}^{p} a_i \boldsymbol{v}_i = \sum_{i=1}^{p} b_i \boldsymbol{v}_i$$

with scalars $a_1, b_1, \ldots, a_p, b_p$ implies $a_i = b_i$ for all $i = 1, \ldots, p$. In other words, the representation of a vector as a linear combination of a finite number of linearly independent vectors is necessarily unique.

As we shall see when discussing spaces of signals such as $L^2(I)$, it will be natural to introduce the following extension of the concept of linear independence: Consider an arbitrary family $\{\boldsymbol{v}_\alpha,\ \alpha \in A\}$ of elements in V with A some index set (not necessarily finite). We say that the vectors $\{\boldsymbol{v}_\alpha,\ \alpha \in A\}$ form a linearly independent family if each of its finite subsets is a linearly independent collection. Formally, this is equivalent to requiring that for every $p = 1, 2, \ldots$ and for every collection $\alpha_1, \ldots, \alpha_p$ of distinct elements in A, the relation

$$\sum_{i=1}^{p} a_i \boldsymbol{v}_{\alpha_i} = \mathbf{0} \tag{3.3}$$

with scalars $a_1, \ldots, a_p$ in $\mathbb{R}$ implies $a_1 = \ldots = a_p = 0$.

3.3 Subspaces and linear spans

A (linear) *subspace* E of the vector space $(V, +)$ (on $\mathbb{R}$) is any subset of V which is closed under vector addition and multiplication by scalars, i.e.,

$$\boldsymbol{v} + \boldsymbol{w} \in E \quad \text{and} \quad a\boldsymbol{v} \in E$$

whenever $\boldsymbol{v}$ and $\boldsymbol{w}$ are elements of E and a is an arbitrary scalar.

Consider an arbitrary family $\{\boldsymbol{v}_\alpha,\ \alpha \in A\}$ of elements in V with A some index set (not necessarily finite). We say that $\boldsymbol{v}$ belongs to the ***(linear) span*** of $\{\boldsymbol{v}_\alpha,\ \alpha \in A\}$, denoted $\mathrm{sp}\,(\boldsymbol{v}_\alpha,\ \alpha \in A)$, if $\boldsymbol{v}$ can be expressed as a linear combination of a *finite* number of elements of $\{\boldsymbol{v}_\alpha,\ \alpha \in A\}$, i.e., there exists a *finite* number of indices in A, say $\alpha_1, \ldots, \alpha_p$ for some p, and scalars $a_1, \ldots, a_p$ in $\mathbb{R}$ such that

$$\boldsymbol{v} = \sum_{i=1}^{p} a_i \boldsymbol{v}_{\alpha_i}.$$

This representation is not a priori unique.

The linear span of this family $\{\boldsymbol{v}_\alpha,\ \alpha \in A\}$ is a linear subspace, and is in fact the smallest linear subspace of E that contains $\{\boldsymbol{v}_\alpha,\ \alpha \in A\}$. In particular, if A is finite, say $A = \{1, \ldots, p\}$ for sake of concreteness, then

$$\mathrm{sp}\,(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p) := \left\{ \sum_{i=1}^{p} a_i \boldsymbol{v}_i : (a_1, \ldots, a_p) \in \mathbb{R}^p \right\}.$$

A subspace E of V is now said to have *dimension* p if there exists p linearly independent vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$ in E (not merely in V) such that $E = \mathrm{sp}\,(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p)$. The notion of dimension is well defined in that if $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_q$ is another collection of linearly independent vectors in E (not merely in V) such that $E = \mathrm{sp}\,(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_q)$, then $p = q$. Any set of p linearly independent vectors $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_p$ such that $E = \mathrm{sp}\,(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_p)$ is called a *basis* of E.

3.4 Scalar product and norm

Many of the vector spaces of interest are endowed with a scalar product, a notion which provides a way to measure correlations between vectors. Formally, a scalar product on the vector space $(V, +)$ is a mapping $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ which satisfies the following conditions

1. (Bilinearity) For each $\boldsymbol{v}$ in V, the mappings $V \to \mathbb{R} : \boldsymbol{w} \to \langle \boldsymbol{v}, \boldsymbol{w} \rangle$ and $V \to \mathbb{R} : \boldsymbol{w} \to \langle \boldsymbol{w}, \boldsymbol{v} \rangle$ are linear mappings, i.e.,

$$\langle \boldsymbol{v}, a\boldsymbol{w} + b\boldsymbol{u} \rangle = a\langle \boldsymbol{v}, \boldsymbol{w} \rangle + b\langle \boldsymbol{v}, \boldsymbol{u} \rangle$$

and

$$\langle a\boldsymbol{w} + b\boldsymbol{u}, \boldsymbol{v} \rangle = a\langle \boldsymbol{w}, \boldsymbol{v} \rangle + b\langle \boldsymbol{u}, \boldsymbol{v} \rangle$$

for all $\boldsymbol{u}$ and $\boldsymbol{w}$ in V, and all scalars a and b in $\mathbb{R}$

2. (Symmetry)
$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{w}, \boldsymbol{v} \rangle, \quad \boldsymbol{v}, \boldsymbol{w} \in V$$

3. (Positive definiteness)
$$\langle \boldsymbol{v}, \boldsymbol{v} \rangle > 0 \quad \text{if} \quad \boldsymbol{v} \neq \mathbf{0} \in V$$

It is easy to see that $\langle \boldsymbol{v}, \boldsymbol{v} \rangle = 0$ when $\boldsymbol{v} = \mathbf{0}$, so that
$$\langle \boldsymbol{v}, \boldsymbol{v} \rangle \geq 0, \quad \boldsymbol{v} \in V.$$

Put differently, $\langle \boldsymbol{v}, \boldsymbol{v} \rangle = 0$ for some vector $\boldsymbol{v}$ in V if and only if $\boldsymbol{v} = \mathbf{0}$.

Once a scalar product is available, it is possible to associate with it a notion of *vector length*. We define a notion of *norm* or vector length on V through the definition

$$\|\boldsymbol{v}\| := \sqrt{\langle \boldsymbol{v}, \boldsymbol{v} \rangle}, \quad \boldsymbol{v} \in V. \tag{3.4}$$

The terminology is justified through the following properties which are commonly associated with the notion of length in Euclidean geometry.

Proposition 3.4.1 *The mapping $V \to \mathbb{R}_+ : \boldsymbol{v} \to \|\boldsymbol{v}\|$ defined by (3.4) satisfies the following properties*

1. *(Homogeneity) For each $\boldsymbol{v}$ in V, it holds that*
$$\|t\boldsymbol{v}\| = |t| \cdot \|\boldsymbol{v}\|, \quad t \in \mathbb{R}.$$

2. *(Positive definiteness) If $\|\boldsymbol{v}\| = 0$ for some $\boldsymbol{v}$ in V, then $\boldsymbol{v} = \mathbf{0}$*

3. *(Triangular inequality) For every pair $\boldsymbol{v}$ and $\boldsymbol{w}$ of elements of V, it holds that*
$$\|\boldsymbol{v} + \boldsymbol{w}\| \leq \|\boldsymbol{v}\| + \|\boldsymbol{w}\|$$

The properties listed in Proposition 3.4.1 form the basis for the notion of norm in more general settings [**?**].

Proof. The homogeneity and positive definiteness are immediate consequence of the definition (3.4) when coupled with the bilinearity of the underlying scalar product and its positive definiteness. To establish the triangular inequality, consider elements $\boldsymbol{v}$ and $\boldsymbol{w}$ of V. It holds that

$$\begin{aligned} \|\boldsymbol{v}+\boldsymbol{w}\|^2 &= \|\boldsymbol{v}\|^2 + \|\boldsymbol{w}\|^2 + 2\langle \boldsymbol{v}, \boldsymbol{w}\rangle \\ &\leq \|\boldsymbol{v}\|^2 + \|\boldsymbol{w}\|^2 + 2\|\boldsymbol{v}\| \cdot \|\boldsymbol{w}\| \\ &= (\|\boldsymbol{v}\| + \|\boldsymbol{w}\|)^2 \qquad (3.5) \end{aligned}$$

where the first equality follows by bilinearity of the scalar product, and the inequality is justified by the Cauchy-Schwartz inequality (discussed in Proposition 3.4.2 below). This establishes the triangular inequality. ■

We conclude this section with a proof of the Cauchy-Schwartz inequality.

Proposition 3.4.2 *The Cauchy-Schwartz inequality*

$$|\langle \boldsymbol{v}, \boldsymbol{w}\rangle| \leq \|\boldsymbol{v}\| \cdot \|\boldsymbol{w}\|, \quad \boldsymbol{v}, \boldsymbol{w} \in V \qquad (3.6)$$

holds with equality in (3.6) if and only if $\boldsymbol{v}$ and $\boldsymbol{w}$ are co-linear, i.e., there exists a scalar a in $\mathbb{R}$ such that $\boldsymbol{v} = a\boldsymbol{w}$.

Proof. Fix $\boldsymbol{v}$ and $\boldsymbol{w}$ elements of V, and note that

$$\begin{aligned} Q(t) &:= \|\boldsymbol{v} + t\boldsymbol{w}\|^2 \\ &= \|\boldsymbol{v}\|^2 + 2t\langle \boldsymbol{v}, \boldsymbol{w}\rangle + t^2\|\boldsymbol{w}\|^2, \quad t \in \mathbb{R} \qquad (3.7) \end{aligned}$$

by bilinearity of the scalar product. The fact that $Q(t) \geq 0$ for *all* t in $\mathbb{R}$ is equivalent to the quadratic equation $Q(t) = 0$ having at most one (double) real root. This forces the corresponding discriminant Δ to be non-positive, i.e.,

$$\Delta = (2\langle \boldsymbol{v}, \boldsymbol{w}\rangle)^2 - 4\|\boldsymbol{v}\|^2\|\boldsymbol{w}\|^2 \leq 0,$$

and the proof of (3.6) is completed. Equality occurs in (3.6) if and only if $\Delta = 0$, in which case there exists $t^\star$ in $\mathbb{R}$ such that $Q(t^\star) = 0$, whence $\boldsymbol{v} + t^\star\boldsymbol{w} = \boldsymbol{0}$, and the co-linearity of $\boldsymbol{v}$ and $\boldsymbol{w}$ follows. ■

In the remainder of this chapter, all discussions are carried out in the context of a vector space $(V, +)$ on $\mathbb{R}$ equipped with a scalar product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$.

3.5 Orthogonality

The elements $\boldsymbol{v}$ and $\boldsymbol{w}$ of V are said to be *orthogonal* if

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle = 0.$$

We also say that the vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ are (pairwise) orthogonal if

$$\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = 0, \quad i \neq j,\ i, j = 1, \ldots, p.$$

More generally, consider an arbitrary family $\{\boldsymbol{v}_\alpha,\ \alpha \in A\}$ of elements in V with A some index set (not necessarily finite). We say that this family is an *orthogonal family* if every one of its finite subset is itself a collection of orthogonal vectors. A moment of reflection shows that this is equivalent to requiring the pairwise conditions

$$\langle \boldsymbol{v}_\alpha, \boldsymbol{v}_\beta \rangle = 0, \quad \alpha \neq \beta \in A. \tag{3.8}$$

Moreover, for any subset E of V, the element $\boldsymbol{v}$ of V is said to be orthogonal to E if

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle = 0, \quad \boldsymbol{w} \in E.$$

If the set E coincides with the linear span of the vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$, then $\boldsymbol{v}$ is orthogonal to E if and only if $\langle \boldsymbol{v}, \boldsymbol{v}_i \rangle = 0$ for all $i = 1, \ldots, p$.

An important consequence of orthogonality is the following version of Pythagoras Theorem.

Proposition 3.5.1 *When $\boldsymbol{v}$ and $\boldsymbol{w}$ are orthogonal elements in V, we have Pythagoras' relation*

$$\|\boldsymbol{v} + \boldsymbol{w}\|^2 = \|\boldsymbol{v}\|^2 + \|\boldsymbol{w}\|^2. \tag{3.9}$$

This result can be used to show a relationship between linear independence and orthogonality.

Lemma 3.5.1 *If the non-zero vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ are orthogonal, then they are necessarily linearly independent.*

Proof. Indeed, for any scalars $a_1, \ldots, a_p$ in $\mathbb{R}$, repeated application of Pythagoras' Theorem yields

$$\|\sum_{i=1}^p a_i \boldsymbol{v}_i\|^2 = \sum_{i=1}^p |a_i|^2 \|\boldsymbol{v}_i\|^2.$$

Therefore, the constraint $\sum_{i=1}^p a_i \boldsymbol{v}_i = 0$ implies $|a_i|^2 \|\boldsymbol{v}_i\|^2 = 0$ for all $i = 1. \ldots, p$. The vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ being non-zero, we have $\|\boldsymbol{v}_i\|^2 \neq 0$ for all $i = 1. \ldots, p$, so that $|a_i|^2 = 0$ for all $i = 1. \ldots, p$. In short, $a_1 = \ldots = a_p = 0$! Thus, the vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ are indeed linearly independent. ■

The notions of orthogonality and norm come together through the notion of orthonormality: If the vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ are orthogonal with unit norm, they are said to be *orthornormal*, a property characterized by

$$\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = \delta(i,j), \quad i,j = 1, \ldots, p. \tag{3.10}$$

The usefulness of this notion is already apparent when considering the following representation result.

Lemma 3.5.2 *If E is a linear space of V spanned by the orthornormal family $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$, then the representation*

$$\boldsymbol{h} = \sum_{i=1}^p \langle \boldsymbol{h}, \boldsymbol{u}_i \rangle \boldsymbol{u}_i, \quad \boldsymbol{h} \in E \tag{3.11}$$

holds, and E has dimension p.

The assumption of Lemma 3.5.2 can always be achieved as should be clear from the Gram-Schmidt orthonormalization procedure discussed in Section 3.8.

Proof. By the definition of E as a span of the vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$, every element $\boldsymbol{h}$ in E is of the form

$$\boldsymbol{h} = \sum_{i=1}^p h_i \boldsymbol{u}_i \tag{3.12}$$

for an appropriate selection of scalars $h_1, \ldots, h_p$. For each $j = 1, \ldots, p$, we find

$$\langle \boldsymbol{h}, \boldsymbol{u}_j \rangle = \langle \sum_{i=1}^p h_i \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = \sum_{i=1}^p h_i \langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = h_j$$

upon invoking orthonormality, and (3.11) follows from (3.12). ■

We emphasize that the discussion of Sections 3.4 and 3.5 depends only on the defining properties of the scalar product. This continues to be the case in the material of the next section.

3.6 Distance and projection

We can define a notion of *distance* on V by setting

$$d(\boldsymbol{v},\boldsymbol{w}) := \|\boldsymbol{v}-\boldsymbol{w}\|, \quad \boldsymbol{v},\boldsymbol{w} \in V. \tag{3.13}$$

Consider now the situation where E is a linear subspace of V and $\boldsymbol{v}$ is an element in V. We are interested in finding an element $\boldsymbol{v}^\star$ in E which has the smallest distance to $\boldsymbol{v}$, namely

$$d(\boldsymbol{v},\boldsymbol{v}^\star) = \inf_{\boldsymbol{x}\in E} d(\boldsymbol{v},\boldsymbol{x}). \tag{3.14}$$

The uniqueness and characterization of such an element $\boldsymbol{v}^\star$ (when it exists) are addressed in

Proposition 3.6.1 *Let E be a linear subspace of V, and let $\boldsymbol{v}$ denote an arbitrary element in V. If there exists an element $\boldsymbol{v}^\star$ in E satisfying (3.14), it is unique and characterized by the simultaneous validity of the relations*

$$\langle \boldsymbol{v}-\boldsymbol{v}^\star, \boldsymbol{h}\rangle = 0, \quad \boldsymbol{h} \in E. \tag{3.15}$$

Conversely, any element $\boldsymbol{v}^\star$ in E satisfying (3.15) necessarily satisfies (3.14).

Before giving the proof of Proposition 3.6.1 in the next section we discuss some easy consequences of the conditions (3.15). These conditions state that the vector $\boldsymbol{v}-\boldsymbol{v}^\star$ is *orthogonal* to E. The unique element $\boldsymbol{v}^\star$ satisfying these constraints is often called the *projection* of $\boldsymbol{v}$ onto E, and at times we shall use the notation

$$\boldsymbol{v}^\star = \mathrm{Proj}_E(\boldsymbol{v}),$$

in which case (3.15) takes the form

$$\langle \boldsymbol{v}-\mathrm{Proj}_E(\boldsymbol{v}), \boldsymbol{h}\rangle = 0, \quad \boldsymbol{h} \in E. \tag{3.16}$$

It is often useful to view $\boldsymbol{v}^\star$ as the *best approximation* of $\boldsymbol{v}$ in E, with $\boldsymbol{v}-\boldsymbol{v}^\star$ interpreted as the *error* incurred by approximating $\boldsymbol{v}$ by $\boldsymbol{v}^\star$. In this interpretation, (3.15) states that the error is orthogonal to the space of all admissible approximations (i.e., those in E). If $\boldsymbol{v}$ is itself an element of E, then $\boldsymbol{v}-\boldsymbol{v}^\star$ is now an element of E and (3.15) (with $\boldsymbol{h}=\boldsymbol{v}-\boldsymbol{v}^\star$ now in E) yields $\|\boldsymbol{v}-\boldsymbol{v}^\star\| = 0$ or equivalently, $\mathrm{Proj}_E(\boldsymbol{v}) = \boldsymbol{v}$, as expected.

For any element $\boldsymbol{v}$ in V whose projection onto E exists, Pythagoras Theorem gives

$$\|\boldsymbol{v}\|^2 = \|\mathrm{Proj}_E(\boldsymbol{v})\|^2 + \|\boldsymbol{v} - \mathrm{Proj}_E(\boldsymbol{v})\|^2 \tag{3.17}$$

as a direct consequence of (3.16)

The linearity of the projection operator is a simple consequence of Proposition 3.6.1 and is left as an exercise to the reader:

Corollary 3.6.1 *For any linear space E of V, the projection mapping $\mathrm{Proj}_E : V \to E$ is a linear mapping wherever defined: For every $\boldsymbol{v}$ and $\boldsymbol{w}$ in V whose projections $\mathrm{Proj}_E(\boldsymbol{v})$ and $\mathrm{Proj}_E(\boldsymbol{w})$ onto E exist, the projection of $a\boldsymbol{v} + b\boldsymbol{w}$ onto E exists for arbitrary scalars a and b in $\mathbb{R}$, and is given by*

$$\mathrm{Proj}_E(a\boldsymbol{v} + b\boldsymbol{w}) = a\mathrm{Proj}_E(\boldsymbol{v}) + b\mathrm{Proj}_E(\boldsymbol{w}).$$

We stress again that at this level of generality, there is no guarantee that the projection always exists. There is however a situation of great practical importance where this is indeed the case.

Lemma 3.6.1 *Assume E to be a linear subspace of V spanned by the orthornormal family $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$ for some finite integer p. Then, every element $\boldsymbol{v}$ in V admits a projection onto E given by*

$$\mathrm{Proj}_E(\boldsymbol{v}) = \sum_{i=1}^{p} \langle \boldsymbol{v}, \boldsymbol{u}_i \rangle \boldsymbol{u}_i. \tag{3.18}$$

For future use, under the conditions of Lemma 3.6.1, we note that

$$\|\mathrm{Proj}_E(\boldsymbol{v})\|^2 = \sum_{i=1}^{p} |\langle \boldsymbol{v}, \boldsymbol{u}_i \rangle|^2, \quad \boldsymbol{v} \in V \tag{3.19}$$

as a simple consequence of the orthonormality of the family $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$.

Proof. Pick an element $\boldsymbol{v}$ in V , and set

$$\boldsymbol{v}^\star := \sum_{i=1}^{p} \langle \boldsymbol{v}, \boldsymbol{u}_i \rangle \boldsymbol{u}_i.$$

The element $\boldsymbol{v}^\star$ belongs to E, with

$$\begin{aligned} \langle \boldsymbol{v} - \boldsymbol{v}^\star, \boldsymbol{u}_i \rangle &= \langle \boldsymbol{v}, \boldsymbol{u}_i \rangle - \langle \boldsymbol{v}^\star, \boldsymbol{u}_i \rangle \\ &= \langle \boldsymbol{v}, \boldsymbol{u}_i \rangle - \sum_{j=1}^{p} \langle \boldsymbol{v}, \boldsymbol{u}_j \rangle \langle \boldsymbol{u}_j, \boldsymbol{u}_i \rangle \\ &= \langle \boldsymbol{v}, \boldsymbol{u}_i \rangle - \langle \boldsymbol{v}, \boldsymbol{u}_i \rangle = 0, \quad i = 1, \ldots, p. \end{aligned} \tag{3.20}$$

From Lemma 3.5.2 it is plain that $\boldsymbol{v} - \boldsymbol{v}^\star$ is orthogonal to E, thus $\boldsymbol{v}^\star$ satisfies (3.15) and the proof is now completed by invoking Proposition 3.6.1. ■

3.7 A proof of Proposition 3.6.1

First, there can be at most one element in E which satisfies (3.15) for if there were two such elements, say $\boldsymbol{v}_1^\star$ and $\boldsymbol{v}_2^\star$ in E, then

$$\langle \boldsymbol{v} - \boldsymbol{v}_k^\star, \boldsymbol{h} \rangle = 0, \qquad \begin{matrix} k = 1, 2 \\ \boldsymbol{h} \in E \end{matrix}$$

so that

$$\langle \boldsymbol{v}_1^\star - \boldsymbol{v}_2^\star, \boldsymbol{h} \rangle = 0, \quad \boldsymbol{h} \in E.$$

Using $\boldsymbol{h} = \boldsymbol{v}_1^\star - \boldsymbol{v}_2^\star$, element of E, in this last relation we find $\|\boldsymbol{v}_1^\star - \boldsymbol{v}_2^\star\| = 0$, whence $\boldsymbol{v}_1^\star = \boldsymbol{v}_2^\star$ necessarily.

Let $\boldsymbol{v}^\star$ be an element in E which satisfies (3.14). For any $\boldsymbol{h}$ in E, the vector $\boldsymbol{v}^\star + t\boldsymbol{h}$ is also an element of E for all t in $\mathbb{R}$. Thus, by the definition of $\boldsymbol{v}^\star$ it holds that

$$\|\boldsymbol{v} - \boldsymbol{v}^\star\|^2 \leq \|\boldsymbol{v} - (\boldsymbol{v}^\star + t\boldsymbol{h})\|^2, \quad t \in \mathbb{R}$$

with

$$\|\boldsymbol{v} - (\boldsymbol{v}^\star + t\boldsymbol{h})\|^2 = \|\boldsymbol{v} - \boldsymbol{v}^\star\|^2 + t^2\|\boldsymbol{h}\|^2 - 2t\langle \boldsymbol{v} - \boldsymbol{v}^\star, \boldsymbol{h} \rangle.$$

Consequently,

$$t^2\|\boldsymbol{h}\|^2 - 2t\langle \boldsymbol{v} - \boldsymbol{v}^\star, \boldsymbol{h} \rangle \geq 0, \quad t \in \mathbb{R}.$$

This last inequality readily implies

$$t\|\boldsymbol{h}\|^2 \geq 2\langle \boldsymbol{v} - \boldsymbol{v}^\star, \boldsymbol{h} \rangle, \quad t > 0$$

and

$$-|t|\|\boldsymbol{h}\|^2 \leq 2\langle \boldsymbol{v} - \boldsymbol{v}^\star, \boldsymbol{h}\rangle, \quad t < 0.$$

Letting t go to zero in each of these last two inequalities yields $\langle \boldsymbol{v} - \boldsymbol{v}^\star, \boldsymbol{h}\rangle \leq 0$ and $\langle \boldsymbol{v} - \boldsymbol{v}^\star, \boldsymbol{h}\rangle \geq 0$, respectively, and the desired conclusion (3.15) follows.

Conversely, consider any element $\boldsymbol{v}^\star$ in E satsifying (3.15). For each $\boldsymbol{x}$ in E, (3.15) implies the orthogonality of $\boldsymbol{v} - \boldsymbol{v}^\star$ and $\boldsymbol{h} = \boldsymbol{v}^\star - \boldsymbol{x}$ (this last vector being in E), and Pythagoras Theorem thus yields

$$\|\boldsymbol{v} - \boldsymbol{x}\|^2 = \|\boldsymbol{v} - \boldsymbol{v}^\star\|^2 + \|\boldsymbol{v}^\star - \boldsymbol{x}\|^2 \geq \|\boldsymbol{v} - \boldsymbol{v}^\star\|^2.$$

This establishes the minimum distance requirement for $\boldsymbol{v}^\star$ and (3.15) indeed characterizes the solution to (3.14). ■

3.8 Gram-Schmidt orthonormalization

As the discussion in Section 3.6 already indicates, the ability to identify $\mathrm{Proj}_E(\boldsymbol{v})$ is greatly simplified if E is spanned by a finite orthonormal family. While E may not be first introduced as being generated by a family of orthonormal vectors, it is however possible to find another family of vectors, this time orthonormal, that nevertheless spans E. The procedure to do so is known as the Gram-Schmidt orthonormalization procedure.

More formally, this procedure provides an algorithm to solve the following problem: Given non-zero vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ in V, find a collection of orthonormal vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{v}_p$ in V such that

$$\mathrm{sp}\left(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\right) = \mathrm{sp}\left(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p\right).$$

While there is no a priori constraint on n, it is plain from previous remarks that $p \leq n$. The Gram-Schmidt procedure is iterative and works as follows:

Step 1: Pick $\boldsymbol{v}_1$ and define the vector $\boldsymbol{u}_1$ by

$$\boldsymbol{u}_1 := \frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|}.$$

This definition is well posed since $\|\boldsymbol{v}_1\| \neq 0$ for the non-zero vector $\boldsymbol{v}_1$. Obviously, $\|\boldsymbol{u}_1\| = 1$. Set

$$\ell(1) := 1 \quad \text{and} \quad E_1 := \mathrm{sp}(\boldsymbol{u}_1),$$

and go to Step 2.

At Step k, the procedure has already returned the ℓ orthonormal vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_\ell$ with $\ell = \ell(k) \leq k$, and let E_ℓ denote the corresponding linear span, i.e., $E_\ell := \mathrm{sp}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_\ell)$.

Step $k+1$**:** Pick $\boldsymbol{v}_{k+1}$.

Either $\boldsymbol{v}_{k+1}$ lies in the span E_ℓ, i.e.,

$$\boldsymbol{v}_{k+1} = \sum_{j=1}^{\ell} \langle \boldsymbol{v}_{k+1}, \boldsymbol{u}_j \rangle \boldsymbol{u}_j,$$

in which case, set

$$\ell(k+1) := \ell(k) \quad \text{and} \quad E_{\ell(k+1)} := E_{\ell(k)}$$

and go to Step $k+2$;

Or $\boldsymbol{v}_{k+1}$ does not lie in E_ℓ, i.e.,

$$\boldsymbol{v}_{k+1} \neq \sum_{j=1}^{\ell} \langle \boldsymbol{v}_{k+1}, \boldsymbol{u}_j \rangle \boldsymbol{u}_j = \mathrm{Proj}_{E_\ell}(\boldsymbol{v}_{k+1}),$$

in which case define

$$\boldsymbol{u}_{\ell+1} := \frac{\boldsymbol{v}'_{k+1}}{\|\boldsymbol{v}'_{k+1}\|}$$

with

$$\begin{aligned} \boldsymbol{v}'_{k+1} &:= \boldsymbol{v}_{k+1} - \mathrm{Proj}_{E_\ell}(\boldsymbol{v}_{k+1}) \\ &= \boldsymbol{v}_{k+1} - \sum_{j=1}^{\ell} \langle \boldsymbol{v}_{k+1}, \boldsymbol{u}_j \rangle \boldsymbol{u}_j. \end{aligned}$$

The algorithm is well defined since $\boldsymbol{v}'_{k+1} \neq \boldsymbol{0}$, while $\boldsymbol{v}'_{k+1}$ is orthogonal to E_ℓ by virtue of (3.16). It is now plain that the vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_\ell, \boldsymbol{u}_{\ell+1}$ form an orthonormal family. Set

$$\ell(k+1) = \ell(k) + 1 \quad \text{and} \quad E_{\ell(k+1)} := \mathrm{sp}\left(E_{\ell(k)} \cup \{\boldsymbol{u}_{\ell(k)+1}\}\right)$$

and go to Step $k+2$.

This algorithm terminates in a finite number of steps, in fact no more than n steps. All the projections encountered in the course of running the algorithm do exist by virtue of Lemma 3.6.1 as they are onto subspaces spanned by a finite number of orthonormal vectors.

3.8.1 Exercises

Ex. 3.1 Show that in a commutative group $(V, +)$, there can be only one zero vector.

Ex. 3.2 Show that in a commutative group $(V, +)$, for every vector $\boldsymbol{v}$ in V, its negative $-\boldsymbol{v}$ is unique.

Ex. 3.3 Let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_q$ denote two collections of linearly independent vectors in V. Show that if $\mathrm{sp}\,(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p) = \mathrm{sp}\,(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_q)$, then necessarily $p = q$.

Ex. 3.4 If E is a linear subspace of V, then it necessarily contains the zero element $\mathbf{0}$. Moreover, $\boldsymbol{v}$ belongs to E if and only if $-\boldsymbol{v}$ belongs to E.

Ex. 3.5 For non-zero vetrors $\boldsymbol{v}$ and $\boldsymbol{w}$ in V, we define their correlation coefficient by

$$\rho(\boldsymbol{v}; \boldsymbol{w}) = \frac{\langle \boldsymbol{v}, \boldsymbol{w} \rangle}{\|\boldsymbol{v}\| \|\boldsymbol{w}\|}.$$

Ex. 3.6 Show that $|\rho(\boldsymbol{v}; \boldsymbol{w})| \leq 1$. Find a necessary and sufficient condition for $\rho(\boldsymbol{v}; \boldsymbol{w}) = 1$ and $\rho(\boldsymbol{v}; \boldsymbol{w}) = -1$.

Ex. 3.7 If the set E is the linear span of the vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ in V, then show that $\boldsymbol{v}$ is orthogonal to E if and only if $\langle \boldsymbol{v}, \boldsymbol{v}_i \rangle = 0$ for all $i = 1, \ldots, p$.

Ex. 3.8 Consider a linear subspece E which is is spanned by the set F in V. Show that $\boldsymbol{v}$ in V is orthogonal to E if and only if $\boldsymbol{v}$is orthogonal to F.

Ex. 3.9 Let E_1 and E_2 be subsets of V such that $E_1 \subseteq E_2$. Assume that for some $\boldsymbol{v}$ in V, its projection $\mathrm{Proj}_{E_2}(\boldsymbol{v})$ exists and is an element of E_1. Explain why

$$\mathrm{Proj}_{E_1}(\boldsymbol{v}) = \mathrm{Proj}_{E_2}(\boldsymbol{v}).$$

Ex. 3.10 Prove Corollary 3.6.1.

Ex. 3.11 Repeat Exercise 3.3 using the Gram-Schmidt orthonormalization procedure.

Ex. 3.12 Let $(V_1, +)$ and $(V_2, +)$ denote two vector spaces on $\mathbb{R}$. A mapping $T : V_1 \to V_2$ is linear if

$$T(a\boldsymbol{v} + b\boldsymbol{w}) = aT(\boldsymbol{v}) + bT(\boldsymbol{w}), \quad \boldsymbol{v}, \boldsymbol{w} \in V_1,\ a, b \in \mathbb{R}.$$

For any subset E of V_1, we write $T(E) = \{T(\boldsymbol{v}),\ \boldsymbol{v} \in E\}$. For E a linear subspace of V_1, show that $T(E)$ is a linear subspace of V_2.

Ex. 3.13 For $i = 1, 2$, let $(V_i, +)$ denote a vector space on $\mathbb{R}$, equipped with its own scalar product $\langle \cdot, \cdot \rangle_i : V_i \times V_i \to \mathbb{R}$, and let $\| \cdot \|_i$ denote the corresponding norm. A mapping $T : V_1 \to V_2$ is said to be norm-preserving if

$$\|T(\boldsymbol{v})\|_2 = \|\boldsymbol{v}\|_1, \quad \boldsymbol{v} \in V_1.$$

Show that if the mapping T is linear, then it is norm-preserving if and only if T preserves the scalar product, i.e.,

$$\langle T(\boldsymbol{v}), T(\boldsymbol{w}) \rangle_2 = \langle \boldsymbol{v}, \boldsymbol{w} \rangle_1, \quad \boldsymbol{v}, \boldsymbol{w} \in V_1.$$

Ex. 3.14

Chapter 4

Finite-dimensional representations

Building on the discussion of Chapter 3, we now present two vector spaces of interest for subsequent developments.

4.1 Finite-dimensional spaces

The simplest example of a vector space is the space $\mathbb{R}^d$ with d some positive integer. An element $\boldsymbol{v}$ of $\mathbb{R}^d$ is identified with the d-uple $(v_1, \ldots, v_d)$ with v_i in $\mathbb{R}$ for each $i = 1, \ldots, d$.

In $\mathbb{R}^d$, the addition and multiplication operations are defined componentwise in the usual way by

$$\boldsymbol{v} + \boldsymbol{w} := (v_1 + w_1, \ldots, v_d + w_d)$$

and

$$a\boldsymbol{v} := (av_1, \ldots, av_d), \quad a \in \mathbb{R}$$

for any pair of vectors $\boldsymbol{v} = (v_1, \ldots, v_d)$ and $\boldsymbol{w} = (w_1, \ldots, w_d)$ in $\mathbb{R}^d$. It is a simple matter to show that these operations turn $(\mathbb{R}^d, +)$ into a vector space on $\mathbb{R}$. The zero element in $(\mathbb{R}^d, +)$ is simply the vector $\mathbf{0} = (0, \ldots, 0)$ with all zero entries.

Statements on the linear independence of vectors in $\mathbb{R}^d$ are statements in Linear Algebra. Indeed, consider vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ in $\mathbb{R}^d$ with $\boldsymbol{v}_i = (v_{i1}, \ldots, v_{id})$ for each $i = 1, \ldots, p$. The linear independentce requirements (3.1) and (3.2) now read as requiring that the d *simultaneous* relations

$$\sum_{i=1}^{p} a_i v_{ij} = 0, \quad j = 1, \ldots, d$$

with scalars $a_1, \ldots, a_p$ in $\mathbb{R}$ imply $a_1 = \ldots = a_p = 0$. In other words, the linear independence of the vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p$ is tantamount to a rank property of the $p \times d$ matrix $\boldsymbol{V} = (v_{ij})$.

The vector space $\mathbb{R}^d$ is endowed with the scalar product given by

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle := \sum_{i=1}^{d} v_i w_i, \quad \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^d.$$

It is a straightforward to check the requisite bilinearity, symmetry and positive definiteness. The norm induced by this scalar product now takes the form

$$\|\boldsymbol{v}\| := \sqrt{(\boldsymbol{v}, \boldsymbol{v})} = \left(\sum_{i=1}^{d} |v_i|^2 \right)^{\frac{1}{2}}, \quad \boldsymbol{v} \in \mathbb{R}^d$$

and the corresponding distance is simply the Euclidean distance on $\mathbb{R}^d$ given by

$$d(\boldsymbol{v}, \boldsymbol{w}) := \|\boldsymbol{v} - \boldsymbol{w}\| = \left(\sum_{i=1}^{d} |v_i - w_i|^2 \right)^{\frac{1}{2}}, \quad \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^d.$$

The vector space $\mathbb{R}^d$ contains a very special set of vectors, denoted by $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_d$, which form an extremely convenient orthonormal family: For each $i = 1, \ldots, d$, the vector $\boldsymbol{e}_i = (e_{i1}, \ldots, e_{id})$ has all its components zero except the i^{th} which is equal to one, i.e.,

$$e_{ij} = \delta(i, j), \quad i, j = 1, \ldots, d.$$

Obviously,

$$\langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle = \delta(i, j), \quad i, j = 1, \ldots, d$$

and for every element $\boldsymbol{v} = (v_1, \ldots, v_d)$ in $\mathbb{R}^d$, we can write

$$\begin{aligned} \boldsymbol{v} &= (v_1, \ldots, v_d) \\ &= v_1(1, 0, \ldots, 0) + v_2(0, 1, \ldots, 0) + v_d(0, 0, \ldots, 1) \\ &= v_1 \boldsymbol{e}_1 + \ldots + v_d \boldsymbol{e}_d. \end{aligned}$$

Thus, $\mathbb{R}^d$ (as a subspace of itself) has dimension d, and therefore no more than d non-zero vectors can ever be orthogonal, hence orthonormal, in $\mathbb{R}^d$.

As an immediate consequence, any linear subspace E of $\mathbb{R}^d$ can always be viewed as the linear span of a *finite* number of orthonormal vectors. Hence, by Lemma 3.6.1 the projection operator onto E is well defined as a mapping $\mathrm{Proj}_E : \mathbb{R}^d \to E$ on the whole of $\mathbb{R}^d$ where it is linear by Corollary 3.6.1.

4.2 Signal spaces

Let I be a non-degenerate interval of the real line $\mathbb{R}$, say $[a,b]$ (with $a < b$), $(-\infty, b]$ or $[a, \infty)$. A (real-valued) signal is any function $\varphi : I \to \mathbb{R}$. The energy of the signal φ is the quantity $\mathcal{E}(\varphi)$ defined by

$$\mathcal{E}(\varphi) := \int_I |\varphi(t)|^2 dt.$$

The signal φ has finite energy if $\mathcal{E}(\varphi) < \infty$. The space of all finite energy signals on the interval I is denoted by $L_2(I)$, namely

$$L_2(I) := \{\varphi : I \to \mathbb{R} : \ \mathcal{E}(\varphi) < \infty\}.$$

The set $L_2(I)$ can be endowed with a vector space structure by introducing a vector addition and multiplication by constants, i.e., for any φ and ψ in $L_2(I)$ and any scalar a in $\mathbb{R}$, we define the signals $\varphi + \psi$ and $a\varphi$ by

$$(\varphi + \psi)(t) := \varphi(t) + \psi(t), \quad t \in I$$

and

$$(a\varphi)(t) := a\varphi(t), \quad t \in I.$$

The signals $\varphi + \psi$ and $a\varphi$ are all finite energy signals if φ and ψ are in $L_2(I)$. It is easy to show that equipped with these operations, $(L_2(I), +)$ is a vector space on $\mathbb{R}$. The zero element for $(L_2(I), +)$ will be the zero signal $\vartheta : I \to \mathbb{R}$ defined by $\vartheta(t) = 0$ for all t in I.

In $L_2(I)$ the notion of linear independence specializes as follows: The signals $\varphi_1, \ldots, \varphi_p$ in $L_2(I)$ are linearly independent if

$$\sum_{i=1}^{p} a_i \varphi_i = \vartheta$$

with scalars $a_1, \ldots, a_p$ in $\mathbb{R}$ implies $a_1 = \ldots = a_p = 0$. This equivalent to the validity of the simultaneous relations

$$\sum_{i=1}^{p} a_i \varphi_i(t) = 0, \quad t \in I$$

with scalars $a_1, \ldots, a_p$ in $\mathbb{R}$ implying $a_1 = \ldots = a_p = 0$. In contrast with the situation in $\mathbb{R}^d$, here there is *no* constraint on p as the following example shows [Exercise 4.7].

Example 4.2.1 *Take $I = [0,1]$ and for each $k = 0, 1, \ldots$, define the signal $\varphi_k : [0,1] \to \mathbb{R}$ by $\varphi_k(t) = t^k$ ($t \in I$). For each $p = 1, 2, \ldots$, the signals $\varphi_0, \varphi_1, \ldots, \varphi_p$ are linearly independent in $L^2(I)$. Therefore, $L_2(I)$ cannot be of finite dimension.*

Here as well, we can define a product scalar by setting

$$\langle \varphi, \psi \rangle := \int_I \varphi(t)\psi(t)dt, \quad \varphi, \psi \in L_2(I).$$

We leave it as an exercise to show that this definition gives rise to a scalar product on $L_2(I)$. The norm of a finite energy signal is now defined by

$$\|\varphi\| := \sqrt{(\varphi, \varphi)}, \quad \varphi \in L_2(I)$$

or in extensive form,

$$\|\varphi\| = \left(\int_I |\varphi(t)|^2 dt \right)^{\frac{1}{2}} = \sqrt{\mathcal{E}(\varphi)}, \quad \varphi \in L_2(I).$$

It should be noted that this notion of "energy norm" is not quite a norm on $L_2(I)$ as understood earlier. Indeed, positive definiteness fails here since $\|\varphi\| = 0$ does not necessarily imply $\varphi = \vartheta$ – Just take $\varphi(t) = 1$ for t in $I \cap \mathbb{Q}$ and $\varphi(t) = 0$ for t in $I \cap \mathbb{Q}^c$, in which case $\|\varphi\| = 0$ but $\varphi \neq \vartheta$! This difficulty is overcome by partitioning $L_2(I)$ into *equivalence classes* with signals considered as equivalent if their difference has zero energy, i.e., the two signals ψ and ψ' in $L_2(I)$ are equivalent if $\|\psi - \psi'\|^2 = 0$. It is this collection of equivalence classes that should be endowed with a vector space structure and a notion of scalar product, instead of the collection of all finite energy signals defined on I – Pointers are provided in Exercises 4.3-4.6. This technical point will be not pursued any further as it does not affect the analyses carried out here. Thus, with a slight abuse of notation, we will consider the "scalar product" defined earlier on $L_2(I)$ as a *bona fide* scalar product.

With these definitions, the notions of orthogonality and orthonormality are defined as before. However, while in $\mathbb{R}^d$ there could be no more than d vectors which can ever be orthonormal, this is not the case in $L_2(I)$ [Exercise 4.8].

Example 4.2.2 *Pick $I = [0,1]$ and for each $k = 0, 1, \ldots$ define the signals $\varphi_k : I \to \mathbb{R}$ by*

$$\varphi_0(t) = 1, \ \varphi_k(t) = \sqrt{2}\cos(2\pi k t), \quad t \in I, \ k = 1, 2, \ldots \tag{4.1}$$

For each $p = 1, 2, \ldots$, the signals $\varphi_0, \varphi_1, \ldots, \varphi_p$ are orthonormal in $L^2(I)$.

The notion of distance on $L_2(I)$ associated with the energy norm takes the special form

$$d(\varphi, \psi) := \left(\int_I |\varphi(t) - \psi(t)|^2 dt \right)^{\frac{1}{2}}, \quad \varphi, \psi \in L_2(I).$$

4.3 Projections in $L_2(I)$

As we now explore the notion of projection onto a linear subspace E of $L_2(I)$, we shall see shortly that in sharp contrast with the situation in $\mathbb{R}^d$, existence is not automatic anymore. In other words, for an arbitrary signal ψ in $L_2(I)$, there is no guarantee that there will always be an element $\psi^\star$ in E which has the smallest distance to ψ, i.e.,

(4.2) $$d(\psi, \psi^\star) = \inf_{\varphi \in E} d(\psi, \varphi).$$

Additional assumptions are needed on E for (4.2) to hold for all signals in $L_2(I)$. However, when $\psi^\star$ does exist, it is necessarily unique by virtue of Proposition 3.6.1.

To gain a better understanding as to why the projection onto E may fail to exist, consider the situation where a *countably infinite* family of orthonormal signals $\{\varphi_k,\ k = 1, 2, \ldots\}$ is available. For each $n = 1, 2, \ldots$, let E_n denote the linear span of the n first signals $\varphi_1, \ldots, \varphi_n$. Fix ψ in $L_2(I)$. By Lemma 3.6.1 the projection of ψ onto E_n always exists, and is given by

$$\widehat{\psi}_n := \mathrm{Proj}_{E_n}(\psi) = \sum_{k=1}^{n} \langle \psi, \varphi_k \rangle \varphi_k,$$

and (3.19) yields

$$\|\widehat{\psi}_n\|^2 = \sum_{k=1}^{n} |\langle \psi, \varphi_k \rangle|^2.$$

With the corresponding error defined by

$$\widetilde{\psi}_n := \psi - \widehat{\psi}_n,$$

we find from (3.17) that

$$\|\psi\|^2 = \|\widehat{\psi}_n\|^2 + \|\widetilde{\psi}_n\|^2$$

by the orthogonality condition (3.15).

Combining these observations leads to

$$\|\widetilde{\psi}_n\|^2 = \|\psi\|^2 - \|\widehat{\psi}_n\|^2 = \|\psi\|^2 - \sum_{k=1}^{n} |\langle \psi, \varphi_k \rangle|^2,$$

and the convergence

$$\lim_{n\to\infty} \|\widetilde{\psi}_n\|^2 := \varepsilon(\psi)$$

takes place in a monotonically decreasing manner. Of course, this is consistent with the geometric viewpoint according to which $\widehat{\psi}_n$ is the best approximation of ψ among the elements of E_n. The inclusions $E_n \subset E_{n+1},\ n = 1, 2, \ldots$ imply that the approximations $\{\widehat{\psi}_n,\ n = 1, 2, \ldots\}$ are increasingly accurate, or equivalently, that the magnitude of the error, namely $\|\widetilde{\psi}_n\|$, decreases.

A natural question is to determine the limiting value $\varepsilon(\psi)$. Several cases arise depending on whether $\varepsilon(\psi) > 0$ or $\varepsilon(\psi) = 0$. In the discussion we make use of the easy identity

$$E_\infty := \mathrm{sp}\,(\varphi_k,\ k = 1, 2, \ldots) = \cup_k E_k. \tag{4.3}$$

Case 1 – If ψ belongs to E_∞, then ψ is an element of E_p for some p and $\widehat{\psi}_{p+k} = \psi$ for all $k = 0, 1, \ldots$, whence $\widetilde{\psi}_{p+k} = \vartheta$, and $\varepsilon(\psi) = 0$. Obviously the projection onto E_∞ does exist with $\psi = \mathrm{Proj}_{E_\infty}(\psi)$.

Case 2 – When ψ is *not* an element of E_∞, then ψ is *not* the zero signal ϑ but two distinct scenarios are possible.

Case 2.a – With ψ not in E_∞, if $\varepsilon(\psi) = 0$, then ψ can be approximated ever closely by an element of E_∞ since $\lim_{n\to\infty} \|\psi - \widehat{\psi}_n\|^2 = 0$. It is then customary to say that ψ is an element of the *closure* of E_∞, a fact noted

$$\psi \in \overline{E_\infty} = \overline{\mathrm{sp}(\varphi_k,\ k = 1, 2, \ldots)}.$$

The set $\overline{E_\infty}$ is called the closure of the linear subspace E_∞; it is itself a linear subspace of $L_2(I)$ which could be defined by

$$\overline{E_\infty} := \{\varphi \in L_2(I) :\ \varepsilon(\varphi) = 0\}.$$

However, $\mathrm{Proj}_{E_\infty}(\psi)$ does *not* exist as the following argument by contradiction shows: If the projection $\mathrm{Proj}_{E_\infty}(\psi)$ were indeed to exist, then it would have to be an element of E_∞, say $\widehat{\psi}$. By the definition of E_∞, the signal $\widehat{\psi}$ is an element of E_p for some p and it is a simple matter to check that $\widehat{\psi} = \widehat{\psi}_{p+k}$ for all $k = 0, 1, \ldots$. Consequently, making use of earlier observations, we find

$$\|\psi\|^2 = \|\widehat{\psi}_{k+p}\|^2 + \|\widetilde{\psi}_{k+p}\|^2 = \|\widehat{\psi}\|^2 + \|\widetilde{\psi}_{k+p}\|^2, \quad k = 0, 1, \ldots$$

Letting k go to infinity and using the fact $\varepsilon(\psi) = 0$, we obtain $\|\psi\|^2 = \|\widehat{\psi}\|^2$. It follows from (3.17) that $\|\widetilde{\psi}\| = 0$ since $\|\psi\|^2 = \|\widehat{\psi}\|^2 + \|\widetilde{\psi}\|^2$ (with $\widetilde{\psi} = \psi - \widehat{\psi}$). Therefore, $\widetilde{\psi} = \vartheta$ and $\psi = \widehat{\psi}$. But this implies that ψ was an element of E_∞ and an contradiction ensues.

On the other hand, $\mathrm{Proj}_{\overline{E_\infty}}(\psi)$ does exist and it is customary to represent it formally as an *infinite* series, namely

$$\mathrm{Proj}_{\overline{E_\infty}}(\psi) = \sum_{k=1}^{\infty} \langle \psi, \varphi_k \rangle \varphi_k, \tag{4.4}$$

to capture the intuitive fact that $\mathrm{Proj}_{\overline{E_\infty}}(\psi)$ is the "limiting" signal increasingly approximated by the projection signals $\{\widehat{\psi}_n, \ n = 1, 2, \ldots\}$. Note that here $\psi = \mathrm{Proj}_{\overline{E_\infty}}(\psi)$.

It follows from the discussion above that only finitely many of the coefficients $\{\langle \psi, \varphi_k \rangle, \ k = 1, 2 \ldots\}$ can ever be zero, and some care therefore needs to be exercised in defining this element (4.4) of $L_2(I)$ – Up to now only finite linear combinations have been considered. For our purpose, it suffices to note that for any sequence $\{c_k, \ k = 1, \ldots\}$ of scalars, the infinite series $\sum_{k=1}^\infty c_k \varphi_k$ can be made to represent an element of $L_2(I)$ under the summability condition

$$\sum_{k=1}^{\infty} |c_k|^2 < \infty. \tag{4.5}$$

This can be achieved by showing that the partial sums

$$\sum_{\ell=1}^{k} c_\ell \varphi_\ell, \quad k = 1, 2, \ldots$$

converge in some suitable sense to an element of $L_2(I)$ (which is represented by $\sum_{k=1}^\infty c_k \varphi_k$). We invite the reader to check that indeed

$$\sum_{k=1}^{\infty} |\langle \psi, \varphi_k \rangle|^2 < \infty, \quad \psi \in L_2(I). \tag{4.6}$$

Example 4.3.1 *Continue with the situation in Example 4.2.2, and set*

$$\psi(t) := \sum_{k=1}^{\infty} \frac{1}{k^2} \cos(2\pi kt), \quad t \in I.$$

The signal ψ is a well defined element of $L_2(I)$ with $\varepsilon(\psi) = 0$, and yet ψ is not an element of E_∞.

Case 2.b – With ψ not in E_∞, if $\varepsilon(\psi) > 0$, then ψ cannot be an element of $\overline{E_\infty}$ and therefore cannot be approximated ever so closely by an element in E_∞. Here $\mathrm{Proj}_{E_\infty}(\psi)$ may not exist, but $\mathrm{Proj}_{\overline{E_\infty}}(\psi)$ always does exist with

$$\psi \neq \mathrm{Proj}_{\overline{E_\infty}}(\psi) = \sum_{k=1}^{\infty} \langle \psi, \varphi_k \rangle \varphi_k.$$

We follow up these comments with the following examples.

Example 4.3.2 *Continue with the situation in Example 4.2.2, and take*

$$\psi(t) := \sin(2\pi t), \quad t \in I.$$

Here, $\varepsilon(\psi) > 0$ and the projection of ψ onto E_∞ exists and $\mathrm{Proj}_{E_\infty}(\psi) = \vartheta$.

Example 4.3.3 *Continue with the situation in Example 4.2.2, and take*

$$\psi(t) := \sin(2\pi t) + \sum_{k=1}^{\infty} \frac{1}{k^2} \cos(2\pi kt), \quad t \in I.$$

This time, it is still the case that $\varepsilon(\psi) > 0$ but the projection of ψ onto E_∞ does not exist.

The last two example show that it is possible to have

$$\overline{E_\infty} \neq L_2(I),$$

a possibility reflecting the fact that the orthonormal family $\{\varphi_k, \ k = 1, 2, \ldots\}$ is not rich enough in that its (finite) linear combinations are not sufficient to approximate some element in $L_2(I)$ to any prescribed level of accuracy. This motivates the following definition: The orthonormal family $\{\varphi_k,\ k = 1, 2, \ldots\}$ is said to be *complete* (in $L_2(I)$) if

$$\overline{E_\infty} = L_2(I).$$

This is equivalent to

$$\varepsilon(\psi) = \lim_{n\to\infty} \|\psi - \widehat{\psi}_n\| = 0$$

for *every* signal ψ in $L_2(I)$.

Example 4.3.4 *Pick $I = [0, 1]$ and for each $k = 0, 1, \ldots$ define the signals $\varphi_k : I \to \mathbb{R}$ by*

$$\varphi_{2k}(t) = \sqrt{2}\cos(2\pi kt), \quad t \in I,\ k = 1, 2, \ldots$$

and

$$\psi_{2k+1}(t) = \sqrt{2}\sin(2\pi kt), \quad t \in I,\ k = 0, 1, \ldots$$

*with $\varphi_0(t) = 1$ $(t \in I)$. It is a non-trivial fact concerning the structure of the space $L_2(I)$ that the orthonormal family $\{\varphi_k,\ k = 0, 1, \ldots\}$ is complete [**?**]*

4.4 Finite-dimensional spaces of $L_2(I)$

The discussion from earlier sections suggests ways to represent finite energy signals. Given an orthonormal family $\{\varphi_k,\ k = 1, 2, \ldots\}$ in $L_2(I)$, we associate with each finite energy signal a sequence of finite dimensional vectors. Formally, for each $n = 1, 2, \ldots$, we set

$$T_n(\psi) := (\langle\psi, \varphi_1\rangle, \ldots, \langle\psi, \varphi_n\rangle), \quad \psi \in L_2(I). \tag{4.7}$$

The vector $T_n(\psi)$ is an element of $\mathbb{R}^n$. By restricting our attention to E_n we get the following useful fact.

Lemma 4.4.1 *For each $n = 1, 2, \ldots$, the correspondence $T_n : E_n \to \mathbb{R}^n$ given by (4.7) is a norm-preserving bijection, i.e., T_n is onto and one-to-one with*

$$\|T_n(\psi)\|^2 = \sum_{k=1}^{n} |\langle\psi, \varphi_k\rangle|^2 = \|\psi\|^2, \quad \psi \in E_n. \tag{4.8}$$

More generally we have

$$\langle T_n(\varphi), T_n(\psi)\rangle = \sum_{k=1}^{n} \langle\varphi, \varphi_k\rangle\langle\psi, \varphi_k\rangle = \langle\varphi, \psi\rangle, \quad \varphi, \psi \in E_n. \tag{4.9}$$

Proof. First, when restricted to E_n, the projection operator Proj_{E_n} reduces to the identity, i.e., $\mathrm{Proj}_{E_n}(\psi) = \psi$ whenever ψ is an element of E_n. Thus, with the notation introduced earlier, for any ψ in E_n, we have

$$\psi = \widehat{\psi}_n = \sum_{k=1}^{n} \langle \psi, \varphi_k \rangle \varphi_k$$

so that

$$\|\psi\|^2 = \sum_{k=1}^{n} |\langle \psi, \varphi_k \rangle|^2$$

and (4.8) holds. The relation (4.9) is proved in a similar way.

As a result, if $T_n(\psi) = T_n(\psi')$ for signals ψ and ψ' in E_n, then $T_n(\psi - \psi') = 0$ by linearity and $\|\psi - \psi'\| = \|T_n(\psi - \psi')\| = 0$ by isometry. The inescapable conclusion is that $\psi = \psi'$, whence T_n is one-to-one.

Finally, any vector $\boldsymbol{v} = (v_1, \ldots, v_n)$ in $\mathbb{R}^n$ gives rise to a signal $\psi_{\boldsymbol{v}}$ in E_n through

$$\psi_{\boldsymbol{v}} := \sum_{k=1}^{n} v_k \varphi_k.$$

It is plain that $\langle \psi_{\boldsymbol{v}}, \varphi_k \rangle = v_k$ for each $k = 1, \ldots, n$, hence $T_n(\psi_{\boldsymbol{v}}) = \boldsymbol{v}$ and the mapping T_n is onto. ■

As a result, any element ψ of E_n can be represented *uniquely* by a vector in $\mathbb{R}^n$. This correspondence, formalized in Lemma 4.4.1, is norm-preserving and allows signals in E_n to be viewed as finite-dimensional vectors.

Next, we address the situation of arbitrary signals. To do so, we will need to assume that the orthonormal family $\{\varphi_k,\ k = 1, 2, \ldots\}$ is rich enough.

Theorem 4.4.1 *Assume the orthonormal family $\{\varphi_k,\ k = 1, 2, \ldots\}$ to be complete in $L_2(I)$. Then, any finite energy signal ψ in $L_2(I)$ admits a unique representation as a sequence*

$$(\langle \psi, \varphi_k \rangle,\ k = 1, 2, \ldots).$$

Moreover, Parseval's identity

$$\|\psi\|^2 = \sum_{k=1}^{\infty} |\langle \psi, \varphi_k \rangle|^2, \quad \psi \in L_2(I) \tag{4.10}$$

holds.

4.5 Exercises

Ex. 4.1 Consider two families $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$ and $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_q$ of linearly independent vectors in $\mathbb{R}^d$. Show that we necessarily have $p = q$ whenever

$$\mathrm{sp}\left(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p\right) = \mathrm{sp}\left(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_q\right).$$

Ex. 4.2 Let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$ be an orthonormal family in $\mathbb{R}^d$ for some integer $p \leq d$. Find the linear span of the family of 2^p vectors in $\mathbb{R}^d$ defined by

$$f(\boldsymbol{b}) := \sum_{\ell=1}^{p} (-1)^{b_\ell + 1} \boldsymbol{u}_\ell$$

with $\boldsymbol{b} = (b_1, \ldots, b_p)$ a binary string of length p, i.e., $b_\ell = 0$ or $b_\ell = 1$ for $\ell = 1, \ldots, p$.

Ex. 4.3 Two signals ψ and ψ' in $L_2(I)$ are said to be equivalent if $\|\psi - \psi'\|^2 = 0$, and we write $\psi \sim \psi'$. Show that this notion defines an equivalence relation on $L_2(I)$.

Ex. 4.4 With the notation of Exercise 4.3, show that addition of signals and multiplication of signals by scalars are compatible with this equivalence relation $\sim$. More precisely, with $\psi \sim \psi'$ and $\varphi \sim \varphi'$ in $L_2(I)$, show that $\psi + \varphi \sim \psi' + \varphi'$ and $a\psi \sim a\psi'$ for every scalar a.

Ex. 4.5 With $\psi \sim \psi'$ and $\varphi \sim \varphi'$ in $L_2(I)$, show that $\|\psi\|^2 = \|\psi'\|^2$ and that $\langle \psi, \varphi \rangle = \langle \psi', \varphi' \rangle$.

Ex. 4.6 Let $\mathcal{L}_2(I)$ denote the collection of equivalence classes induced on $L_2(I)$ by the equivalence relation $\sim$. Using Exercise 4.4 and Exercise 4.5, define a structure of vector space on $\mathcal{L}_2(I)$ and a notion of scalar product.

Ex. 4.7 Show that the signals $\{\varphi_k,\ k = 0, 1, \ldots\}$ of Example 4.2.1 are linearly independent in $L_2(I)$.

Ex. 4.8 Show that the signals $\{\varphi_k,\ k = 0, 1, \ldots\}$ of Example 4.2.2 form an orthonormal family in $L_2(I)$.

Ex. 4.9 Apply the Gram-Schmidt orthonormalization procedure to the family $\{\varphi_k,\ k = 0,1,2\}$ in $L_2[0,1]$ given by

$$\varphi_k(t) = t^k, \qquad \begin{array}{c} t \in [0,1] \\ k = 0,1,2 \end{array}$$

Does the answer depend on the order in which the algorithm processes the signals φ_0, φ_1 and φ_2?

Ex. 4.10 The *distinct* finite energy signals $\psi_1, \ldots, \psi_n$ defined on $[0,1]$ have the property that $\psi_1(t) = \ldots = \psi_n(t)$ for all t in the subinterval $[\alpha, \beta]$ with $0 < \alpha < \beta < 1$. Are such signals necessarily linearly independent in $L_2[0,1]$? Explain.

Ex. 4.11 Starting with a finite energy signal g in $L_2[0,T]$ with $\mathcal{E}(g) > 0$, define the two signals g_c and g_s in $L_2(0,T)$ by

$$g_c(t) := g(t)\cos(2\pi f_c t) \quad \text{and} \quad g_s(t) := g(t)\sin(2\pi f_c t), \quad 0 \le t \le T$$

for some carrier frequency $f_c > 0$. Show that the signals g_c and g_s are always linearly independent in $L_2[0,T]$.

Ex. 4.12 Consider the M signals $s_1, \ldots, s_M$ in $L_2[0,T]$ given by

$$s_m(t) = A\cos(2\pi f_c t + \theta_m), \qquad \begin{array}{c} 0 \le t \le T \\ m = 1, \ldots, M \end{array}$$

with amplitude $A > 0$, carrier $f_c > 0$ and distinct phases $0 \le \theta_1 < \ldots < \theta_M < 2\pi$. What is the dimension L of $\mathrm{sp}\,(s_1, \ldots, s_M)$? Find an orthonormal family in $L_2[0,T]$, say $\varphi_1, \ldots, \varphi_L$, such that $\mathrm{sp}\,(s_1, \ldots, s_M) = \mathrm{sp}\,(\varphi_1, \ldots, \varphi_L)$. Find the corresponding finite dimensional representation.

Ex. 4.13 Apply the Gram-Schmidt orthonormalization procedure to the family of M signals given in Exercise 4.12.

Ex. 4.14 Same problem as in Exercise 4.12. for the M signals given by

$$s_m(t) = A_m g(t), \quad \begin{array}{c} 0 \le t \le T \\ m = 1, \ldots, M \end{array}$$

with g a pulse in $L_2[0,T]$ and distinct amplitudes $A_1 < \ldots < A_M$.

Ex. 4.15 Apply the Gram-Schmidt orthonormalization procedure to the family of M signals given in Exercise 4.14.

Ex. 4.16 For the collection $\{\varphi_k,\ k = 0, 1, \ldots\}$ in Example 4.2.1, find φ in $L_2(0, 1)$ such that φ does not belong to the linear span $\mathrm{sp}(\varphi_k,\ k = 0, 1, \ldots)$, but does belong to its closure $\overline{\mathrm{sp}(\varphi_k,\ k = 0, 1, \ldots)}$.

Ex. 4.17 Consider a set $\{s_1, \ldots, s_M\}$ of M linearly dependent signals in $L_2[0, T)$. Now partition the interval $[0, T)$ into K non-empty subintervals, say $[t_k, t_{k+1})$ $(k = 0, \ldots, K - 1)$ with $t_0 = 0$ and $t_M = T$. For each $k = 1, \ldots, K$, let $\boldsymbol{\alpha}_k = (\alpha_{k1}, \ldots, \alpha_{kM})$ denote an element of $\mathbb{R}^M$, and define the new constellation $\{s_1^\star, \ldots, s_M^\star\}$ by

$$s_m^\star(t) = \alpha_{km} s_m(t), \quad t \in [t_{k-1}, t_k),\ k = 1, \ldots, K$$

for each $m = 1, \ldots, M$. Find conditions on the original constellation $\{s_1, \ldots, s_M\}$ and on the vectors $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_K$ that ensure the linear independence (in $L_2(0, T)$) of the signals $\{s_1^\star, \ldots, s_K^\star\}$.

Ex. 4.18 Consider a finite energy non-constant pulse $g : [0, 1] \to \mathbb{R}$, with $g(t) > 0$ in the unit interval $[0, 1]$. Are the signals g and g^2 linearly independent in $L_2[0, 1]$? Are the signals $g, g^2, \ldots, g^p$ always linearly independent in $L_2[0, 1]$?

Ex. 4.19 For each $\alpha > 0$, let s_α and c_α denote the signals $\mathbb{R} \to \mathbb{R}$ given by

$$s_\alpha(t) = \sin(\alpha t) \quad \text{and} \quad c_\alpha(t) = \cos(\alpha t), \quad t \in \mathbb{R}.$$

For $T > 0$ and $\alpha \neq \beta$, find conditions for each of the collections $\{s_\alpha, c_\alpha\}$, $\{s_\alpha, s_\beta\}$, $\{s_\alpha, c_\alpha, s_\beta\}$ and $\{s_\alpha, c_\alpha, s_\beta, c_\beta\}$ (restricted to the interval $[0, T]$) to be orthogonal in $L_2(0, T)$.

Ex. 4.20 Show (4.3).

Ex. 4.21 Discuss Example 4.3.1.

Ex. 4.22 Discuss Example 4.3.2.

Ex. 4.23 Discuss Example 4.3.3.

Ex. 4.24