2 MATHEMATICAL BACKGROUND

We will assume that readers have basic knowledge of linear algebra and multivariable calculus. Since these skills may be a little rusty, however, here we recall an incomplete set of key ideas that will be used throughout our discussion. Readers comfortable with these topics can skip this chapter, although it is recommended they review especially the multivariable calculus sections and the discussion of Einstein notation.

2.1 LINEAR ALGEBRA

Although its simplicity and elegance place linear algebra in the pantheon of basic mathematical tools, linear algebra developed much later than geometry. While Euclid's *Elements* formalized aspects of geometry circa 300 BC, systematic study of linear problems did not take off until the seventeenth century [2]; that said, versions of basic linear algebra techniques such as Gaussian elimination appeared for millennia without being codified. At this point, however, we recognize linear algebra as a foundational tool both computationally and theoretically. Underscoring this point, the term "matrix" in linear algebra was invented in 1850 by J. Sylvester from the Latin word for *womb* [60].

2.1.1 *Basic Definitions*

The main object of study in linear algebra is the *vector space*. For our purposes, rather than considering a general field F, we consider vector spaces over the set of real numbers $\mathbb{R}$:

Definition 2.1 (Vector space over $\mathbb{R}$)**.** *A* vector space *over $\mathbb{R}$ is a set V equipped with two binary operations $+$ and $\cdot$ satisfying the following axioms:*

- GROUP AXIOMS: *For $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, we have*

$$\begin{aligned} \mathbf{u} + (\mathbf{v} + \mathbf{w}) &= (\mathbf{u} + \mathbf{v}) + \mathbf{w} && \textit{(associativity)} \\ \mathbf{u} + \mathbf{v} &= \mathbf{v} + \mathbf{u} && \textit{(commutativity).} \end{aligned}$$

  In addition, there exists a zero element *$\mathbf{0} \in V$ such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$, and for each $\mathbf{v} \in V$ there exists an additive inverse $-\mathbf{v}$ such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$.*

- VECTOR SPACE AXIOMS: *For $a, b \in \mathbb{R}$ and $\mathbf{u}, \mathbf{v} \in V$, we have*

$$\begin{aligned} a(\mathbf{u} + \mathbf{v}) &= a\mathbf{u} + a\mathbf{v} && \textit{(distributivity)} \\ (a + b)\mathbf{v} &= a\mathbf{v} + b\mathbf{v} && \textit{(distributivity)} \\ a(b\mathbf{v}) &= (ab)\mathbf{v} && \textit{(associativity)} \\ 1\mathbf{v} &= \mathbf{v} && \textit{(identity).} \end{aligned}$$

Our notation above drops the operator $\cdot$ for multiplication. Elements of V are called vectors*; in this context we label real numbers in $\mathbb{R}$ as* scalars*. A subset of a vector space that is itself a vector space is a* subspace*.*

Occasionally it will be useful to use complex rather than real numbers; the definition above suffices for that case, simply replacing $\mathbb{R}$ with the set of complex numbers $\mathbb{C} := \{a + b\sqrt{-1} : a, b \in \mathbb{R}\}$.

From a high level, a vector space is any set of objects that can be (1) added together and (2) scaled. These two basic operations define a *linear combination* $\mathbf{w} \in V$ of elements $\{\mathbf{v}_k\}_{k=1}^m \subset V$ of a vector space V with weights $\{a^k\}_{k=1}^m$:

$$\mathbf{w} = a^1\mathbf{v}_1 + a^2\mathbf{v}_2 + \cdots + a^m\mathbf{v}_m = \sum_{k=1}^m a^k\mathbf{v}_k.$$

To be clear, in this notation values like a^2 represent the second element of a, rather than the squared value of a scalar a. See the aside in §2.1.3 for discussion of this notation.

The typical example of a vector space is Euclidean space $\mathbb{R}^n$, defined as the set of n-dimensional vectors. It is important to remember, however, that many other vector spaces exist that do not resemble Euclidean space. Two examples are below:

Example 2.1 (Polynomials)**.** *The set of polynomials of degree at most d is a vector space. Recall that a polynomial of degree at most $d \in \mathbf{N}$ is any function $\mathbf{p}(t)$ that can be written*

$$\mathbf{p}(t) = \sum_{k=0}^d a^k\mathbf{p}_k(t),$$

where the a^k's are fixed constants and $\mathbf{p}_k(t) := t^k$. Adding two degree-d polynomials together creates a third, as does scaling a polynomial by a constant:

$$\left(\sum_{k=0}^d a^k\mathbf{p}_k(t)\right) + \left(\sum_{k=0}^d b^k\mathbf{p}_k(t)\right) = \sum_{k=0}^d (a^k + b^k)\mathbf{p}_k(t)$$
$$c\left(\sum_{k=0}^d a^k\mathbf{p}_k(t)\right) = \sum_{k=0}^d (ca^k)\mathbf{p}_k(t).$$

A straightforward exercise is to check that the axioms in Definition 2.1 hold for this set. In contrast, multiplying two degree-d polynomials can make a degree-$2d$ polynomial, leaving the space; thankfully, Definition 2.1 only requires multiplication between scalars and vectors, not between two vectors.

Example 2.2 (L^2 functions)**.** *The set $L^2([0,1])$ is comprised of all functions $f : [0,1] \to \mathbb{R}$ that are (Lebesgue) square-integrable:*

$$\int_0^1 f(t)^2\,dt < \infty.$$

In particular, the integral must be defined and finite; in this case the axioms in Definition 2.1 reflect properties of the integral, e.g. that the integral of a sum of functions is the sum of their integrals.

The smallest subspace of a vector space V containing a set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V$ is the *span*, defined by:

$$\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} := \left\{\mathbf{w} = \sum_{k=1}^m a^k\mathbf{v}_k : \{a^k\}_{k=1}^m \subset \mathbb{R}\right\}.$$

The span of a set of vectors can be thought of as the set of all vectors in V "reachable" by linearly combining vectors in the set.

It could be the case that adding a new vector to a set does not change its span. For example, in $\mathbb{R}^2$ we know $\operatorname{span}\{(1,0)\} = \operatorname{span}\{(1,0),(2,0)\}$. This is because $(2,0) = 2\cdot(1,0)$, that is, $(2,0)$ is already in the span of $\{(1,0)\}$. Hence, the latter case is somehow redundant. This idea is formalized in the definition of linear dependence:

Definition 2.2 (Linear dependence)**.** *A set of vectors $W \subseteq V$ is* linearly dependent *if there exists a vector $\mathbf{w} \in W$ that equals a linear combination of vectors in $W\backslash\{\mathbf{w}\}$. Otherwise the set is* linearly independent*.*

This leads to a key measure of the complexity of a vector space, its *dimension*:

Definition 2.3 (Dimension)**.** *The* dimension *of a vector space V is the maximum number of linearly independent vectors in any subset $W \subseteq V$.*

A set of linearly independent vectors that spans V is called a *basis* for V. Any two bases for V contain the same number of vectors, namely the dimension of V, and every $\mathbf{v} \in V$ can be written as a linear combination of basis vectors in exactly one way.

Example 2.3 (Standard basis for $\mathbb{R}^n$)**.** *The dimension of Euclidean space $\mathbb{R}^n$ is n, spanned by the standard basis vectors*

$$\mathbf{e}_k \coloneqq (\underbrace{0, \dots, 0,}_{k-1 \text{ elements}} \; 1, \; \underbrace{0, \dots, 0}_{n-k \text{ elements}}).$$

Example 2.4 (Infinite-dimensional vector spaces)**.** *Not all vector spaces have a finite dimension. For example, the space $L^2([0,1])$ from Example 2.2 is infinite-dimensional. While we know this intuitively (there are* many *functions!), one way to prove this is to show that the Fourier basis functions $f_k(t) \coloneqq \sin \pi k t$ are linearly independent for all $k \in \mathbf{N}$.*

While we often illustrate vectors by drawing arrows, the reality of Definition 2.1 is that it is not particularly geometric. For example, it does not require there to be a notion of the length of a vector or the distance between two vectors. Formally, geometric structure is provided by *inner products*:

Definition 2.4 (Inner product)**.** *An* inner product *on a vector space V is a map $g(\cdot, \cdot) : V \times V \to \mathbb{R}$ satisfying the following properties:*

$$\begin{aligned} g(\mathbf{u}, \mathbf{v}) &= g(\mathbf{v}, \mathbf{u}) && \textit{(symmetry)} \\ g(a\mathbf{u}, \mathbf{v}) &= a g(\mathbf{u}, \mathbf{v}) && \textit{(linearity)} \\ g(\mathbf{u} + \mathbf{v}, \mathbf{w}) &= g(\mathbf{u}, \mathbf{w}) + g(\mathbf{v}, \mathbf{w}) && \textit{(linearity)} \\ g(\mathbf{u}, \mathbf{u}) &\geq 0 \textit{ with equality only when } \mathbf{u} = \mathbf{0} && \textit{(positive definiteness).} \end{aligned}$$

A key example of an inner product is the dot product $\mathbf{v} \cdot \mathbf{w} \coloneqq \sum_{k=1}^{n} v^k w^k$ on Euclidean space $\mathbb{R}^n$, where v^k and w^k denote the entries of $\mathbf{v}$ and $\mathbf{w}$ (see §2.1.3). A different example of an inner product is the following:

Example 2.5 (Inner product of functions)**.** *Given $f, g \in L^2([0,1])$, the L^2 inner product $\langle f, g \rangle$ is given by $\langle f, g \rangle \coloneqq \int_0^1 f(t) g(t)\, dt$.*

Inner products give a notion of length to a vector, via the definition of a *norm*:

$$\|\mathbf{v}\|_g \coloneqq \sqrt{g(\mathbf{v}, \mathbf{v})}.$$

We will use the notation $\|\mathbf{v}\|_2$ to denote the 2-norm of a vector from the dot product defined above.

Inner products also determine angles between vectors, motivated by the identity

$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\|_2 \|\mathbf{w}\|_2 \cos\theta$$

in $\mathbb{R}^3$ as well as the more general Cauchy–Schwarz inequality

$$|g(\mathbf{v}, \mathbf{w})| \leq \|\mathbf{v}\|_g \|\mathbf{w}\|_g,$$

which suggests that higher-dimensional inner products also have some notion of cosine between -1 and 1.

Using this idea to extend the idea of perpendicular vectors to dimension greater than three leads to the definition of *orthogonal* vectors $\mathbf{v}, \mathbf{w} \in V$ as vectors satisfying $g(\mathbf{v}, \mathbf{w}) = 0$. An *orthonormal basis* is a basis of vectors $\mathbf{v}_1, \mathbf{v}_2 \dots$ spanning a vector space V such that

$$g(\mathbf{v}_k, \mathbf{v}_\ell) = \begin{cases} 1 & \text{if } k = \ell \\ 0 & \text{otherwise.} \end{cases}$$

2.1.2 *Linear Maps*

The power of linear algebra is in its characterization of *linear maps*, or maps from one vector space into another that preserve addition and scalar multiplication:

Definition 2.5 (Linear map)**.** *A map $L[\cdot] : V \to W$ from vector space V into vector space W is* linear *if for any $\mathbf{u}, \mathbf{v} \in V$ and $c \in \mathbb{R}$ we have*

$$\begin{aligned} L[\mathbf{u} + \mathbf{v}] &= L[\mathbf{u}] + L[\mathbf{v}] \quad \textit{(additivity)} \\ L[c\mathbf{u}] &= cL[\mathbf{u}] \quad \textit{(homogeneity)}. \end{aligned}$$

A corollary of this definition is that linear maps preserve the weights of linear combinations:

$$L\left[\sum_{k=1}^{p} a^k \mathbf{v}_k\right] = \sum_{k=1}^{p} a^k L[\mathbf{v}_k],$$

for any collection of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p \in V$ and scalars $a^1, \ldots, a^p \in \mathbb{R}$. Linear maps also necessarily take the zero element of one vector space to the zero element of another.

The best-known linear operators are given as matrix-vector products, explored in §2.1.3, but others exist. One example is as follows:

Example 2.6 (Linear operator)**.** *Consider the following operator from $L^2([0,1])$ into $\mathbb{R}$:*

$$L[f] \coloneqq \int_0^1 f(t)\, dt.$$

This map takes elements of one vector space, functions in $L^2([0,1])$, and maps them to elements of another vector space, scalars in $\mathbb{R}^1$. Checking that L is linear involves two tests:

$$\begin{aligned} L[f+g] &= \int_0^1 [f(t) + g(t)]\, dt = \int_0^1 f(t)\, dt + \int_0^1 g(t)\, dt = L[f] + L[g] \\ L[cf] &= \int_0^1 cf(t)\, dt = c\int_0^1 f(t)\, dt = cL[f]. \end{aligned}$$

The *kernel* or *null space* of a linear operator $L : V \to W$ is the set of vectors L maps to the zero element of W:

$$\operatorname{kernel} L \coloneqq \{\mathbf{v} \in V : L[\mathbf{v}] = \mathbf{0}\}.$$

The *image* of L is the subspace of W corresponding to vectors mapped under L:

$$\operatorname{image} L \coloneqq \{\mathbf{w} \in W : \mathbf{w} = L[\mathbf{v}] \text{ for some } \mathbf{v} \in V\}.$$

In the finite-dimensional case, the image is also known as the *column space*, since it is the span of the columns of L represented as a matrix; the dimension of the column space is known as the *rank*.

We can use an inner product g to make a few measurements about the action of a linear operator $L : V \to V$, which maps a vector space V onto itself. First, take any orthonormal basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of V. The *trace* of L quantifies how much L stretches each basis vector along itself:

$$\operatorname{tr} L \coloneqq \sum_{k=1}^{n} g(\mathbf{v}_k, L[\mathbf{v}_k]).$$

Second, the *Frobenius norm* of L adds up the squared lengths of the images of the basis vectors,

$$\|L\|_{\text{Fro}}^2 \coloneqq \sum_{k=1}^{n} \|L[\mathbf{v}_k]\|_g^2 = \sum_{k=1}^{n} g(L[\mathbf{v}_k]\|, L[\mathbf{v}_k]).$$

A basic (but non-obvious) fact is that neither quantity depends on which orthonormal basis we choose; rather, they are properties of L alone (see exercise 2.2.). We will see what they look like in coordinates in §2.1.3.

2.1.3 *Euclidean Space*

Our discussion above has been generic to any vector space over the real numbers, but most computational applications of linear algebra involve a single vector space, Euclidean space $\mathbb{R}^n$.

A vector in $\mathbb{R}^n$ can be represented using a tuple of values $\mathbf{v} = (v^1, v^2, \ldots, v^n)$, which serves as a compact way of writing $\mathbf{v}$ in the standard basis introduced in Example 2.3:

$$\mathbf{v} = \sum_{k=1}^{n} v^k \mathbf{e}_k.$$

A third more space-consuming option that aligns well with calculations in linear algebra is to write $\mathbf{v}$ as an $n \times 1$ column:

$$\mathbf{v} = \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{pmatrix}. \tag{2.1}$$

All of these coordinate-wise representations in terms of $v^1, \ldots, v^n$ implicitly involve the standard basis vectors $\mathbf{e}_i$, which is of course not the only basis for $\mathbb{R}^n$.

Aside (Einstein notation)**.** *Note our usage of* upper indices v^k *for the coefficients of* $\mathbf{v}$ *in the standard basis, notated with* lower indices $\mathbf{e}_k$. *This choice is not by accident but rather suggestive of a convention in differential geometry known as* Einstein notation, *in which any time an index is repeated twice (once upper, once lower) it is implicitly summed. For instance, we could write:*

$$\mathbf{v} = v^k \mathbf{e}_k.$$

This notation can be confusing to readers who do not work with it every day, so with the possible exception of a few tedious calculations we will avoid dropping summation operators.

Suppose $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear map. Applying the basic properties of linearity, for any vector $\mathbf{v} \in \mathbb{R}^n$ we have

$$L[\mathbf{v}] = L\left[\sum_{k=1}^{n} v^k \mathbf{e}_k\right] = \sum_{k=1}^{n} v^k L[\mathbf{e}_k].$$

In other words, the map L is *completely determined* by its action on the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. This suggests representing L using a *matrix*

$$L \sim \begin{pmatrix} | & | & & | \\ L[\mathbf{e}_1] & L[\mathbf{e}_2] & \cdots & L[\mathbf{e}_n] \\ | & | & & | \end{pmatrix} \in \mathbb{R}^{m \times n}.$$

We repeat here that this representation of L is *in the standard basis*, which may not be canonical for a problem that is rotated from the xyz axes.

There are many advantages of the matrix representation for L. In particular, it makes sense computationally to store a grid of numbers, and a vector is a special case of a matrix with a single column, as in (2.1). In addition, matrix-vector multiplication (and matrix-matrix multiplication as a generalization) admits a simple mnemonic, moving from left to right in the first term and top to bottom in the second:

$$L[\mathbf{v}] = \underbrace{\begin{pmatrix} | & | & & | \\ L[\mathbf{e}_1] & L[\mathbf{e}_2] & \cdots & L[\mathbf{e}_n] \\ | & | & & | \end{pmatrix}}_{\longrightarrow} \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{pmatrix}\Bigg\downarrow = \sum_{k=1}^{n} v^k L[\mathbf{e}_k].$$

For consistency with our (admittedly pedantic) choice of index notation, when a matrix represents a linear map we will use one upper index and one lower index to represent its individual elements in the standard basis:

$$\begin{pmatrix} L_1^1 & L_2^1 & \cdots & L_n^1 \\ L_1^2 & L_2^2 & \cdots & L_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ L_1^m & L_2^m & \cdots & L_n^m \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^n L_k^1 v^k \\ \sum_{k=1}^n L_k^2 v^k \\ \vdots \\ \sum_{k=1}^n L_k^m v^k \end{pmatrix} \coloneqq \begin{pmatrix} w^1 \\ w^2 \\ \vdots \\ w^m \end{pmatrix}. \tag{2.2}$$

Here, L_j^i is the coefficient of $L[\mathbf{e}_j]$ in the $\mathbf{e}_i$ direction: $L_j^i = \mathbf{e}_i \cdot L[\mathbf{e}_j]$. Notice this notation is somewhat misleading when $m \neq n$: $\mathbf{e}_i \in \mathbb{R}^m$ and $\mathbf{e}_j \in \mathbb{R}^n$, so they are technically standard basis vectors in different dimensions.

When $m = n$, the trace and Frobenius norm defined in §2.1.2 take a simple form in the standard basis. Since $\mathbf{e}_i \cdot L[\mathbf{e}_i] = L_i^i$, the trace is the sum of diagonal entries ($\operatorname{tr} L = \sum_i L_i^i$), and the Frobenius norm is the square root of the sum of the squares of all the entries ($\|L\|_{\text{Fro}}^2 = \sum_{ij} (L_j^i)^2$).

Matrices also appear when considering inner products between vectors; we will distinguish this case by changing the positions of the indices. Suppose $g(\cdot, \cdot)$ is an inner product on $\mathbb{R}^n$. By bilinearity, we can write:

$$g(\mathbf{v}, \mathbf{w}) = g\left(\sum_{k=1}^n v^k \mathbf{e}_k, \sum_{\ell=1}^n w^\ell \mathbf{e}_\ell\right) = \sum_{k,\ell=1}^n v^k w^\ell g(\mathbf{e}_k, \mathbf{e}_\ell). \tag{2.3}$$

Hence, if we notate $g_{k\ell} \coloneqq g(\mathbf{e}_k, \mathbf{e}_\ell)$ then in Einstein notation we have $g(\mathbf{v}, \mathbf{w}) = v^k w^\ell g_{k\ell}$.

We can also store the values $g_{k\ell}$ in a matrix, but the entries of the matrix have a different meaning: They are inner products of the standard basis vectors instead of images of the standard basis vectors under a linear map. This distinction explains our different choice of index positions. This aside, if we define $G \in \mathbb{R}^{n \times n}$ as a matrix with entries $g_{k\ell}$, we have

$$g(\mathbf{v}, \mathbf{w}) = \sum_{k,\ell=1}^n g_{k\ell} v^k w^\ell = \mathbf{v}^\top G \mathbf{w},$$

where we think of $\mathbf{v}$ and $\mathbf{w}$ as $n \times 1$ matrices and the result of the product is a 1×1 matrix, or scalar. Here, $\top$ denotes the *transpose* operator, which flips the two indices of a matrix. In particular, the dot product on $\mathbb{R}^n$ can be written compactly as $\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^\top \mathbf{w}$, a formula we will use frequently.

Example 2.7 (Dot product)**.** *From our discussion above, it is evident that the dot product is an inner product with the choice $G = I_{n \times n}$, the identity matrix. In other words, the implicit assumption in computing a dot product is that the inner product of two standard basis vectors is given by*

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} \coloneqq \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

Unless otherwise noted, we will assume $\mathbb{R}^n$ is equipped with the dot product as its inner product.

Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is an orthonormal basis for $\mathbb{R}^n$ under the dot product, other than the standard one. Collect the coordinates of the $\mathbf{v}_k$'s in the standard basis as the columns of a matrix $Q \in \mathbb{R}^{n \times n}$. Then, the entries of $Q^\top Q$ are exactly the dot products $\mathbf{v}_k \cdot \mathbf{v}_\ell$, so orthonormality of the basis is equivalent to writing $Q^\top Q = I_{n \times n}$; a matrix satisfying this condition is called *orthogonal*. Multiplying by Q converts coordinates in the basis $\{\mathbf{v}_k\}$ into coordinates in the standard basis, and multiplying by $Q^\top = Q^{-1}$ goes the other way. Orthogonal matrices preserve dot products, as $(Q\mathbf{v}) \cdot (Q\mathbf{w}) = \mathbf{v}^\top Q^\top Q \mathbf{w} = \mathbf{v} \cdot \mathbf{w}$, and hence lengths and angles: They are exactly the rotations and reflections of $\mathbb{R}^n$.

2.1.4 *Linear Systems of Equations*

Countless contexts in linear algebra—and nearly any other branch of mathematics—require the solution of *inverse problems*, in which we attempt to undo the action of some mathematical operator. A counterintuitive theme is that solving a *forward problem* and finding its inverse require different machinery. For example, multiplying a set of prime numbers to find a product is easy even when the product is a fairly large number, while finding the prime factors of a large integer can take tons of computational time. While inverting a linear map is not nearly as challenging as prime factorization, the inversion of a linear map L is a formidable algorithmic task to carry out efficiently and stably.

Given a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$ represented as a matrix $A \in \mathbb{R}^{m\times n}$ and a vector $\mathbf{b} \in \mathbb{R}^m$, the basic linear system of equations asks to find $\mathbf{x} \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{b}$. This problem is solvable any time $\mathbf{b}$ is in the column space of A (regardless of m and n!).

Basic algorithms for solving $A\mathbf{x} = \mathbf{b}$ often assume A is *invertible*, meaning one can find exactly one $\mathbf{x}$ corresponding to any $\mathbf{b}$. The linear map $\mathbf{x} \mapsto A\mathbf{x}$ certainly must be bijective for A to be invertible, so by dimension-counting we must have $m = n$ in this case. Accompanying the condition that A is square, several equivalent conditions imply invertibility of A:

- The columns (or rows) of A span $\mathbb{R}^n$.
- The columns (or rows) of A are linearly independent.
- $A\mathbf{x} = \mathbf{0}$ implies $\mathbf{x} = \mathbf{0}$.
- The determinant of A is nonzero.

In any of these cases, A admits an inverse matrix A^{-1} so that $AA^{-1} = A^{-1}A = I_{n\times n}$.

A considerable portion of introductory linear algebra coursework is dedicated to solving $A\mathbf{x} = \mathbf{b}$ by hand, but the reality is that there exist relatively stable and efficient algorithms and software for doing this in practice. Typical algorithms include Gaussian elimination, LU factorization, and iterative methods like conjugate gradients (CG). In our discussion, we will assume that a "black box" implementation of a linear solver is at our disposal, so we do not need to discuss these algorithms in detail; see [26] for detailed discussion.

One notable algorithmic consideration for solving linear systems is the identification of *structure* in the matrix A. If we know more information about the contents of A, we may be able to choose more efficient algorithms than the sledgehammer techniques used to invert any $A \in \mathbb{R}^{n\times n}$. Properties that can be leveraged include the following:

- *Positive semidefinite* matrices A (notated $A \succeq 0$) are symmetric matrices satisfying $\mathbf{x}^\top A\mathbf{x} \geq 0$ for any $\mathbf{x} \in \mathbb{R}^n$; the *positive definite* matrices additionally satisfy the property that $\mathbf{x}^\top A\mathbf{x} = 0$ if and only if $\mathbf{x} = \mathbf{0}$. In this case, we can always factor $A = LL^\top$ where L is lower-triangular, using the Cholesky factorization. Then, $A^{-1} = L^{-\top}L^{-1}$, where the inverse of L and its transpose are algorithmically easier to compute.
- The entries of *sparse* matrices are mostly zero. This is by far the most common scenario in low-dimensional geometry processing and in graph theory, since often each row/column of a matrix corresponds to a single vertex on a mesh with nonzeros only when vertices share an edge. Sparse matrices might be stored using a long list of triplets $(\text{row}, \text{column}, \text{value})$ rather than as an $m \times n$ grid of mostly zero values. Specialized algorithms factor and invert sparse matrices in time roughly proportional to the number of nonzero elements rather than mn.
- *Structured* matrices might have repeated blocks, which may be easier to store once rather than copying the blocks multiple times.

- Some matrices are easy to apply but hard to write explicitly. For instance, for fixed matrices $A, B \in \mathbb{R}^{n\times n}$ the following is a linear operator:

$$f(C) = CA + BC.$$

  We could try to recover C from $f(C)$ by writing C as a long vector in $\mathbb{R}^{n^2}$ and then figuring out the corresponding $n^2 \times n^2$ matrix representing the action of f, but this calculation is tedious and inefficient. Instead, it may be easier to provide a piece of code that implements the linear map $C \mapsto f(C)$ without giving the elements of the corresponding matrix explicitly, in which case iterative algorithms like gradient descent and conjugate gradients can be preferable for inverting f.

The guide above is vague in the sense that we do not provide algorithmic details. This is on purpose: It is rare to need to implement a numerical linear algebra algorithm in practice, since well-tested implementations are available for free. Languages such as Python and Matlab have these built-in, and libraries such as Eigen [30] and SuiteSparse [20] provide implementations in languages like C++.

Warning. *There is rarely a need to compute A^{-1} from $A \in \mathbb{R}^{n\times n}$ explicitly. Many algorithms can solve the system $A\mathbf{x} = \mathbf{b}$ from a single $\mathbf{b}$ much faster than computing all of A^{-1} and then multiplying to obtain $\mathbf{x} = A^{-1}\mathbf{b}$. Furthermore, inverting matrices explicitly can ruin their structure; for instance, the inverse of a sparse matrix is rarely sparse.*

Example 2.8 (Least-squares)**.** *Suppose we wish to solve the following least-squares problem for $\mathbf{x} \in \mathbb{R}^n$ given $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$:*

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|A\mathbf{x} - \mathbf{b}\|_2^2. \tag{2.4}$$

Intuition for this problem is illustrated in Figure REF*. This problem appears when we are not confident if $\mathbf{b}$ is in the column space of A, and hence we want to find the $\mathbf{x}$ for which $A\mathbf{x}$ is as close as possible to $\mathbf{b}$. Using the techniques e.g. in §2.2.2, we can show that the optimal $\mathbf{x}$ satisfies the* normal equations

$$A^\top A\mathbf{x} = A^\top \mathbf{b}. \tag{2.5}$$

We can check that $A^\top A$ is positive semidefinite:

$$\mathbf{x}^\top A^\top A\mathbf{x} = (A\mathbf{x})^\top (A\mathbf{x}) = \|A\mathbf{x}\|_2^2 \geq 0.$$

The matrix $A^\top A$ is invertible when A has linearly independent columns. If A is square then in this case A is invertible. If A is non-square but has linearly independent columns, the least-squares problem is overdetermined *but still (2.5) determines a unique solution for (2.4). When the columns of A are linearly dependent, $A^\top A$ is not invertible, indicating (2.4) has a nonunique solution.*

2.1.5 *Eigenproblems*

The last linear algebra problem we consider is the *eigenvector* problem:

Definition 2.6 (Eigenvector)**.** *An* eigenvector *of a linear transformation $L : V \to V$ acting on vector space V is a vector $\mathbf{v} \in V\backslash\{\mathbf{0}\}$ such that*

$$L[\mathbf{v}] = \lambda\mathbf{v} \tag{2.6}$$

for some $\lambda \in \mathbb{R}$. In this case, λ is known as the eigenvalue *corresponding to eigenvector $\mathbf{v}$.*

As illustrated in Figure REF, geometrically eigenvectors are directions in which the linear transformation L scales a vector without rotating it.

Example 2.9 (Eigenvector of a matrix)**.** *The eigenvectors of the matrix*

$$M := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

are given by

$$\mathbf{v}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \quad \textit{and} \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

In particular,

$$M\mathbf{v}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = -1 \cdot \mathbf{v}_1$$
$$M\mathbf{v}_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 \cdot \mathbf{v}_2.$$

The corresponding eigenvalues of M are $\lambda_1 = -1$ and $\lambda_2 = 1$.

Example 2.9 illustrates a slight abuse of language typical in the linear algebra world. Here, "the eigenvectors" of M include not only $\mathbf{v}_1$ and $\mathbf{v}_2$ but also these two vectors scaled by any nonzero real number.

Example 2.10 (Eigenfunctions)**.** *Take $C^\infty([0,1])$ to be the set of smooth functions $f : [0,1] \to \mathbb{R}$. Define the operator $L : C^\infty([0,1]) \to C^\infty([0,1])$ given by the second derivative*

$$L[f] = f''.$$

For instance,

$$L[t^2] = \frac{d^2}{dt^2} t^2 = 2 \quad \textit{and} \quad L[\log(t+1)] = \frac{d^2}{dt^2} \log(t+1) = -\frac{1}{(t+1)^2}.$$

Then, eigenfunctions of L include e^{at} and $\cos at$, since

$$L[e^{at}] = a^2 \cdot e^{at} \quad \textit{and} \quad L[\cos at] = -a^2 \cdot \cos at.$$

The corresponding eigenvalues are a^2 and $-a^2$.

We conclude our discussion of classic linear algebra by quoting a famous theorem regarding the eigenvectors of a symmetric matrix:

Theorem 2.1 (Spectral theorem)**.** *If $A \in \mathbb{R}^{n \times n}$ is symmetric, there exists an orthonormal basis for $\mathbb{R}^n$ (with respect to the dot product $\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^\top \mathbf{w}$) of eigenvectors of A.*

In matrix language, if we collect the eigenvectors of a symmetric matrix A as the columns of an orthogonal matrix Q and the corresponding eigenvalues on the diagonal of a diagonal matrix Λ, the spectral theorem states that we can factor $A = Q\Lambda Q^\top$. It also links eigenvalues to definiteness: Writing $\mathbf{x} = Q\mathbf{y}$, we have $\mathbf{x}^\top A\mathbf{x} = \mathbf{y}^\top \Lambda \mathbf{y} = \sum_k \lambda_k (y^k)^2$, so a symmetric matrix is positive semidefinite exactly when all of its eigenvalues are nonnegative and positive definite exactly when they are all positive.

2.2 MULTIVARIABLE CALCULUS

To consider *curved* rather than flat geometry, we need multivariable tools that extend beyond the reach of linear algebra. That is, most spaces are nonlinear—although we may approximate them locally with linear objects. Multivariable calculus equips us with the machinery we need to understand maps in the more general case.

2.2.1 *Optimization*

A key motivator for differential calculus is to study *optimization* problems. In one variable, the connection between optimization and calculus is derived by noticing that—under a differentiability assumption—any minimum of a function $f(x) : \mathbb{R} \to \mathbb{R}$ must be a root of the first derivative

$f'(x)$. Hence, we can search for minima among the roots of f'; the sign of the second derivative $f''(x)$ then determines whether critical points $x \in \mathbb{R}$ satisfying $f'(x) = 0$ are minima, maxima, or indeterminate. We follow a similar path here for functions on $\mathbb{R}^n$ to define the gradient and Hessian.

2.2.2 *Gradients and Hessians*

Suppose $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ is a function[1] that assigns a real value to every point in $\mathbb{R}^n$. In analogy to one-dimensional optimization problems, we might wish to find an $\mathbf{x} \in \mathbb{R}^n$ that achieves the minimum value of f locally or globally.

As illustrated in Figure REF, if a point $\mathbf{x}_0 \in \mathbb{R}^n$ is a local minimum of f, then in particular it locally minimizes f along any one-dimensional slice through $\mathbf{x}_0$. Formalizing this idea, take an arbitrary $\mathbf{v} \in \mathbb{R}^n$ and define the *differential* of f at $\mathbf{x}_0$ as

$$df_{\mathbf{x}_0}(\mathbf{v}) := \lim_{h \to 0} \frac{f(\mathbf{x}_0 + h\mathbf{v}) - f(\mathbf{x}_0)}{h}.$$

As a motivator for proofs we will do in differential geometry, we note the following property:

Proposition 2.1. *$df_{\mathbf{x}_0} : \mathbb{R}^n \to \mathbb{R}$ is a linear operator.*

Proof. Applying one-dimensional calculus n times shows

$$f(\mathbf{x}_0 + h\mathbf{v}) = f(\mathbf{x}_0) + h \sum_{k=1}^{n} \frac{\partial f}{\partial x^k} v^k + O(h^2), \tag{2.7}$$

where we expand $\mathbf{v} = \sum_k v^k \mathbf{e}_k$. Here, $\frac{\partial f}{\partial x^k}$ denotes the k-th *partial derivative*, or derivative of f considered as a one-variable function along coordinate k. The notation $O(h^2)$ is informal shorthand for "terms no larger than a constant times h^2, which become negligible relative to the terms we have written out as $h \to 0$. Plugging this expansion into the definition of $df_{\mathbf{x}_0}$ yields the desired result. □

While the previous proposition does not require *any* inner product on $\mathbb{R}^n$ to make sense, a closer examination of (2.7) (or application of exercise 2.4.) shows existence of a vector $\nabla f(\mathbf{x}_0) \in \mathbb{R}^n$ so that

$$df_{\mathbf{x}_0}(\mathbf{v}) = \nabla f(\mathbf{x}_0) \cdot \mathbf{v}.$$

Note the *direction* of ∇f might change if we equip $\mathbb{R}^n$ with a different inner product; see exercise 2.4. for details.

In any event, we define a *critical point* of a function $f : \mathbb{R}^n \to \mathbb{R}$ to be any $\mathbf{x}_0 \in \mathbb{R}^n$ satisfying $df_{\mathbf{x}_0} \equiv 0$, or equivalently $\nabla f(\mathbf{x}_0) = \mathbf{0}$. If f is differentiable, any local minimum or maximum of f is a critical point. If we wish to minimize f, after finding its critical points we have to verify which of the critical points are local minima. Continuing to apply the logic of Figure REF, $\mathbf{x}_0 \in \mathbb{R}^n$ is a local minimum of f only if it is a local minimum along any line through $\mathbf{x}_0$.

A similar argument to that of Proposition 2.1 shows existence of a *bilinear* operator $H_{\mathbf{x}_0}(\mathbf{v}, \mathbf{w})$ such that $H_{\mathbf{x}_0}(\mathbf{v}, \mathbf{v}) = \frac{d^2}{dh^2} f(\mathbf{x}_0 + h\mathbf{v})\Big|_{h=0}$; in Euclidean coordinates this justifies the Taylor expansion

$$f(\mathbf{x}_0 + \mathbf{v}) = f(\mathbf{x}_0) + \mathbf{v} \cdot \nabla f(\mathbf{x}_0) + \frac{1}{2} \mathbf{v}^\top H_{\mathbf{x}_0} \mathbf{v} + O(\|\mathbf{v}\|_2^3),$$

where $H_{\mathbf{x}_0}$ in this context is the *Hessian matrix* of second partial derivatives of f; when $H_{\mathbf{x}_0}$ is positive definite at a critical point $\mathbf{x}_0$, we know $\mathbf{x}_0$ is a local minimum of f.

1 Here and in our subsequent discussion, we will ignore (important) considerations regarding differentiability and instead will assume all functions are smooth.

Example 2.11 (Least-squares). *Suppose we wish to minimize the function*

$$f(\mathbf{x}) = \frac{1}{2}\|A\mathbf{x} - \mathbf{b}\|_2^2,$$

where $A \in \mathbb{R}^{m\times n}$ *and* $\mathbf{b} \in \mathbb{R}^m$. *The gradient of* f *is*

$$\nabla_{\mathbf{x}} f(\mathbf{x}) = A^\top A\mathbf{x} - A^\top \mathbf{b}.$$

Setting this expression to zero verifies the famous normal equations

$$A^\top A\mathbf{x} = A^\top \mathbf{b}$$

characterizing the solution to least-squares problems. The Hessian of f *has no dependence on* $\mathbf{x}$*:*

$$H \equiv A^\top A.$$

$A^\top A$ *is positive definite when* A *has full column rank, verifying that* $\mathbf{x} = (A^\top A)^{-1}A^\top \mathbf{b}$ *is the unique global minimizer of the least-squares problem in this case.*

We also can extend the differential to maps with more than one output. Such maps will be our main tool for describing curves and surfaces: A parameterized curve, for instance, is a map from $\mathbb{R}$ into $\mathbb{R}^3$.

Suppose $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^m$ has m outputs $f^1(\mathbf{x}), \ldots, f^m(\mathbf{x})$, each a function $\mathbb{R}^n \to \mathbb{R}$. Nothing in the definition of the differential depended on f having a single output, so we can define

$$df_{\mathbf{x}_0}(\mathbf{v}) := \lim_{h\to 0} \frac{f(\mathbf{x}_0 + h\mathbf{v}) - f(\mathbf{x}_0)}{h} \in \mathbb{R}^m,$$

which simply stacks the differentials of the outputs f^i into a vector. By Proposition 2.1 applied to each output, $df_{\mathbf{x}_0}$ is a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$, so by our discussion in §2.1.3 it can be written as an $m \times n$ matrix. This matrix is the *Jacobian* $Df(\mathbf{x}_0)$ of f at $\mathbf{x}_0$, whose entries are partial derivatives:

$$(Df(\mathbf{x}_0))^i_j := \frac{\partial f^i}{\partial x^j}(\mathbf{x}_0), \qquad \text{so that} \qquad df_{\mathbf{x}_0}(\mathbf{v}) = Df(\mathbf{x}_0)\mathbf{v}.$$

Row i of the Jacobian is the gradient $\nabla f^i(\mathbf{x}_0)$ written as a row vector, and when $m = 1$ the Jacobian equals $\nabla f(\mathbf{x}_0)^\top$. As in (2.2), the Jacobian carries one upper and one lower index, since it represents a linear map. It is the best linear approximation of f near $\mathbf{x}_0$, in the sense that $f(\mathbf{x}_0 + \mathbf{v}) = f(\mathbf{x}_0) + Df(\mathbf{x}_0)\mathbf{v} + O(\|\mathbf{v}\|_2^2)$.

Jacobians compose similarly to the one-dimensional case. For $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}^p$, the *chain rule* states

$$D(g \circ f)(\mathbf{x}_0) = Dg(f(\mathbf{x}_0))\, Df(\mathbf{x}_0),$$

the product of a $p \times m$ matrix and an $m \times n$ matrix. In words, the linear approximation of a composition of maps is the composition of their linear approximations.

2.2.3 *Lagrange Multipliers and KKT Conditions*

A unique aspect of multivariable calculus is that it is possible to add *constraints* to an optimization problem. That is, rather than minimizing a function $f(\mathbf{x})$ over all possible $\mathbf{x} \in \mathbb{R}^n$, we might restrict admissible $\mathbf{x}$ values to some subset: $\mathbf{x} \in S \subseteq \mathbb{R}^n$; a point satisfying the constraint is called *feasible*. While constraints are easy to deal with when $n = 1$ since they typically just specify a range of $\mathbf{x}$ values, e.g. $x \in [a, b]$, as shown in Figure REF higher-dimensional constraints might specify curves or regions.

Suppose in addition to our objective function $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ we are given a second function $g(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^m$. As illustrated in Figure REF, we could ask to find the minimum of a function $f(\mathbf{x})$ subject to the constraint that $\mathbf{x}$ is in the *zero level set* $\{\mathbf{x} : g(\mathbf{x}) = 0\}$:

$$\begin{array}{ll} \min_{\mathbf{x}\in\mathbb{R}^n} & f(\mathbf{x}) \\ \text{subject to} & g(\mathbf{x}) = \mathbf{0}. \end{array} \tag{2.8}$$

The function $g(\mathbf{x})$ takes values in $\mathbb{R}^m$, implicitly encoding m potentially independent constraints $g_1(\mathbf{x}) = 0, \ldots, g_m(\mathbf{x}) = 0$.

Our goal is to derive first-order conditions for local optima for the constrained problem, similar to $\nabla f = \mathbf{0}$ in the unconstrained case. To give intuition for these conditions, we take the coarsest possible approximation of f and the components of g:

$$\begin{aligned} f(\mathbf{x}_0 + \mathbf{v}) &= f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^\top \mathbf{v} + O(\|\mathbf{v}\|_2^2) \\ g_k(\mathbf{x}_0 + \mathbf{v}) &= g_k(\mathbf{x}_0) + \nabla g_k(\mathbf{x}_0)^\top \mathbf{v} + O(\|\mathbf{v}\|_2^2)\ \forall k \in \{1, \ldots, m\}. \end{aligned}$$

Suppose $\mathbf{x}_0$ satisfies the constraints: $g(\mathbf{x}_0) = \mathbf{0}$. By the expressions above, a small perturbation $\mathbf{v}$ of $\mathbf{x}_0$ maintaining the constraint must satisfy $\nabla g_k(\mathbf{x}_0)^\top \mathbf{v} = 0$ for all k. In other words, we need $\mathbf{v}$ to be orthogonal to the span of the $\nabla g_k(\mathbf{x}_0)$'s:

$$\mathbf{v} \perp \text{span}\,\{\nabla g_k(\mathbf{x}_0) : k \in \{1, \ldots, m\}\} = \{\sum_k \lambda^k \nabla g_k(\mathbf{x}_0) : \lambda^1, \ldots, \lambda^m \in \mathbb{R}\}.$$

On the other hand, a perturbation $\mathbf{v}$ to (strictly) decrease from $f(\mathbf{x}_0)$ to $f(\mathbf{x}_0 + \mathbf{v})$ must satisfy $\nabla f(\mathbf{x}_0)^\top \mathbf{v} \neq 0$.

Putting these arguments together, a local optimum $\mathbf{x}_0$ of f subject to $g = \mathbf{0}$ has the property that *any* small perturbation $\mathbf{v}$ of $\mathbf{x}_0$ that decreases f must violate the constraints. In symbols, there exist $\lambda^1, \ldots, \lambda^m$ such that

$$\nabla f(\mathbf{x}_0) = -\sum_{k=1}^m \lambda^k \nabla g_k(\mathbf{x}_0). \tag{2.9}$$

The sign on the right-hand side is arbitrary and chosen for a convention we will establish below. The coefficients λ^k are known as *Lagrange multipliers* or *dual variables*, providing a first-order optimality condition for equality-constrained optimization. Intuition for this condition is illustrated in Figure REF.

If we define the *Lagrangian* function

$$\Lambda(\mathbf{x}; \lambda) \coloneqq f(\mathbf{x}) + \sum_{k=1}^m \lambda^k g_k(\mathbf{x}),$$

then it is easy to check that the Lagrange multiplier condition (2.9) is equivalent to finding a critical point of Λ with respect to both $\mathbf{x} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}^m$ simultaneously. One tricky aspect of this observation is explored in problem 2.5.: Solutions of the constrained optimization problem (2.8) are minima with respect to $\mathbf{x}$ and maxima with respect to λ.

Example 2.12 (Eigenvalue problems)**.** *For symmetric, positive definite $A \in \mathbb{R}^{n\times n}$, consider the optimization problem*

$$\begin{array}{ll} \min_{\mathbf{x}\in\mathbb{R}^n} & \mathbf{x}^\top A\mathbf{x} \\ \textit{subject to} & \|\mathbf{x}\|_2^2 = 1. \end{array}$$

This problem has Lagrangian

$$\Lambda(\mathbf{x}; \lambda) = \mathbf{x}^\top A\mathbf{x} + \lambda(1 - \|\mathbf{x}\|_2^2).$$

Setting the derivative with respect to $\mathbf{x}$ to zero shows

$$A\mathbf{x} = \lambda\mathbf{x}.$$

This gives an optimization-oriented interpretation of the eigenvector of a positive definite matrix with the smallest eigenvalue, as the minimizer of $\mathbf{x}^\top A\mathbf{x}$ with a unit norm constraint; note without the constraint the minimizer would simply be $\mathbf{x} = \mathbf{0}$.

As illustrated in Figure REF, a more general constraint can be expressed as $h(\mathbf{x}) \leq \mathbf{0}$, where $h(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^p$; here, we use the convention that $h(\mathbf{x}) \leq \mathbf{0} \iff h_1(\mathbf{x}) \leq 0, \dots, h_p(\mathbf{x}) \leq 0$. This is a generalization of the equality constraints discussed above, since $h(\mathbf{x}) \leq 0$ and $-h(\mathbf{x}) \leq 0$ together imply $h(\mathbf{x}) = 0$.

Suppose $\mathbf{x}_0 \in \mathbb{R}^n$ is a local optimum for the problem

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & h(\mathbf{x}) \leq \mathbf{0}. \end{aligned} \tag{2.10}$$

Then, for each component h_k, $k \in \{1, \dots, p\}$, we have two possibilities:

- h_k is *active* at $\mathbf{x}_0$ if $h_k(\mathbf{x}_0) = 0$. Typically (although not always!), if a constraint is active it implies that removing the constraint from the optimization problem would affect the position of the optimal $\mathbf{x}_0$. In other words, minimizing f without constraint k would result in an even lower value of f. Conversely, if we changed constraint k from an inequality constraint to an equality constraint, we would not expect the solution to change.
- h_k is *inactive* at $\mathbf{x}_0$ if $h_k(\mathbf{x}_0) < 0$ strictly. In this case, assuming h_k is continuous we know that $h_k(\mathbf{x}) < 0$ in some neighborhood of $\mathbf{x}_0$, so removing constraint k would not locally affect the position of the local minimum $\mathbf{x}_0$.

We can extend Lagrange multipliers to this case. By the logic above, we could throw away all inactive constraints at a critical point $\mathbf{x}_0$ and convert active inequality constraints to equality constraints $h_k(\mathbf{x}_0) = 0$. Define λ^k to be either 0 if a constraint is inactive or to be the Lagrange multiplier for the equality-constrained problem at a critical point. We can write this definition compactly via the "complementary slackness" constraint:

$$0 = \lambda^k h_k(\mathbf{x}_0) \; \forall k \in \{1, \dots, p\}.^2 \tag{2.11}$$

This constraint on λ^k forces it to be zero if h_k is nonzero. Additionally, since we zeroed out irrelevant terms we still have the "stationarity condition"

$$0 = \nabla f(\mathbf{x}_0) + \sum_{k=1}^{p} \lambda^k \nabla h_k(\mathbf{x}_0). \tag{2.12}$$

It is acceptable if we perturb $\mathbf{x}_0$ in a direction that converts an active constraint into an inactive one, that is, if we move to the *interior* of constraint k. From (2.9), we can think of the nonzero λ^k's as the components of $-\nabla f(\mathbf{x}_0)$ in the $\nabla h_k(\mathbf{x}_0)$ basis. Recalling that $-\nabla f(\mathbf{x}_0)$ is the first-order *descent* direction for f, we can enforce that any perturbation that decreases f will violate an active constraint via the *dual feasibility* constraint

$$\lambda^k \geq 0. \tag{2.13}$$

In total, conditions (2.11)–(2.13) together with the "primal feasibility" constraint $h(\mathbf{x}_0) \leq \mathbf{0}$ form the Karush–Kuhn–Tucker (KKT) conditions for inequality-constrained optimization.

Example 2.13 (Nonnegative projection)**.** *Suppose we wish to find the closest vector $\mathbf{x} \in \mathbb{R}^n$ to a vector $\mathbf{x}_0 \in \mathbb{R}^n$ with only nonnegative elements. We can express this as an optimization problem:*

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \quad & \|\mathbf{x} - \mathbf{x}_0\|_2^2 \\ \textit{subject to} \quad & \mathbf{x} \geq \mathbf{0}. \end{aligned}$$

2 This is *not* Einstein notation: This product is zero for each k independently.

In this case, we can write $h_k(\mathbf{x}) = -\mathbf{x} \cdot \mathbf{e}_k$, the negative of the k-th component of the vector $\mathbf{x}$. Then, the KKT conditions are as follows:

$$\begin{aligned} \textit{Complementary slackness:} \quad & \lambda^k(\mathbf{x} \cdot \mathbf{e}_k) = 0 \\ \textit{Stationarity:} \quad & 2(\mathbf{x} - \mathbf{x}_0) - \textstyle\sum_{k=1}^n \lambda^k \mathbf{e}_k = 0 \\ \textit{Dual feasibility:} \quad & \lambda^1, \dots, \lambda^n \geq 0 \\ \textit{Primal feasibility:} \quad & \mathbf{x} \cdot \mathbf{e}_1, \dots, \mathbf{x} \cdot \mathbf{e}_n \geq 0. \end{aligned}$$

Taking the dot product of both sides of the stationarity condition with $\mathbf{e}_k$ shows

$$2\mathbf{e}_k \cdot (\mathbf{x} - \mathbf{x}_0) = \lambda^k.$$

We have two cases:

1. *If $\lambda^k = 0$, we know $\mathbf{x} \cdot \mathbf{e}_k = \mathbf{x}_0 \cdot \mathbf{e}_k$, in other words, the k-th entry in $\mathbf{x}$ is identical to that of $\mathbf{x}_0$. By primal feasibility, this can only happen when $\mathbf{x}_0 \cdot \mathbf{e}_k \geq 0$.*
2. *If $\lambda^k \neq 0$, then by complementary slackness we know $\mathbf{x} \cdot \mathbf{e}_k = 0$; in this case $\lambda^k = -2\mathbf{x}_0 \cdot \mathbf{e}_k$, so dual feasibility requires $\mathbf{x}_0 \cdot \mathbf{e}_k \leq 0$.*

Combining these shows $\mathbf{x} = \max(\mathbf{x}_0, \mathbf{0})$, where the max is taken entrywise. In other words, the closest projection of $\mathbf{x}_0$ onto the nonnegative vectors simply zeroes out negative entries.

2.2.4 *Integration*

Our final task is to define a basic notion of an integral from multivariable calculus. Given a continuous function $f : \mathbb{R}^n \to \mathbb{R}$, we might ask for the integral $\int_S f(\mathbf{x})\, d\mathbf{x}$ over some subset $S \subseteq \mathbb{R}^n$. We will not get into the particularities of specifically which sets S are reasonable as domains of integration but rather assume here that S is not a pathological set over which integrals are not defined.

To begin, we might consider what it means to integrate f over a rectangle $R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$. In this case, we can define integration recursively using the *iterated integral*

$$\int_R f(\mathbf{x})\, d\mathbf{x} \coloneqq \int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_n}^{b_n} f(x^1, \dots, x^n)\, dx^n \cdots dx^1.$$

This is actually a recursive set of one-dimensional integrals rather than the integral of a function over a set in $\mathbb{R}^n$.

Given a more general set $S \subseteq \mathbb{R}^n$, define the *indicator function*

$$\chi_S(\mathbf{x}) \coloneqq \begin{cases} 1 & \text{if } \mathbf{x} \in S \\ 0 & \text{otherwise.} \end{cases}$$

Then, we can define the integral of f over S as the integral of the product:

$$\int_S f(\mathbf{x})\, d\mathbf{x} \coloneqq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x^1, \dots, x^n) \chi_S(x^1, \dots, x^n)\, dx^n \cdots dx^1. \tag{2.14}$$

Once again, this iterated integral only requires nested calculations in one-dimension at a time. We will defer integration over more general domains like curves and surfaces until we need it.

2.3 EXERCISES

2.1. Prove the following generalization of the spectral theorem: Suppose $A, B \in \mathbb{R}^{n \times n}$ are symmetric and positive definite. Then, there exists a basis of generalized eigenvectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \subset \mathbb{R}^n$ satisfying $A\mathbf{x}_k = \lambda_k B\mathbf{x}_k$ for some $\lambda_k \in \mathbb{R}$ that are orthogonal under the B inner product $g_B(\mathbf{v}, \mathbf{w}) \coloneqq \mathbf{v}^\top B\mathbf{w}$.

2.2. **JS: Show that the trace and Frobenius norm defined in §2.1.2 do not depend on the choice of orthonormal basis.**

2.3. **JS: Verify some matrix calculus identities.**

2.4. Suppose $\mathbb{R}^n$ is equipped with a general inner product $g(\cdot, \cdot)$. Show that any linear operator $L : \mathbb{R}^n \to \mathbb{R}^1$ can be written $L(\mathbf{v}) = g(\mathbf{v}_0, \mathbf{v})$ for some fixed $\mathbf{v}_0 \in \mathbb{R}^n$ depending on L and g but not $\mathbf{v}$. Use this fact to define a "generalized gradient vector" of $f : \mathbb{R}^n \to \mathbb{R}$ with respect to inner product g, notated $\nabla_g f(\mathbf{x})$.

2.5. **JS: Minimax version of Lagrange/KKT**