Applied Optimization

Ross Baldick

Appendices

Appendix A

Mathematical preliminaries

In this appendix we introduce some notation and mathematical background that we will need throughout the book.

A.1 Notation

A.1.1 Universal and existential quantifiers

Definition A.1 The symbol $\forall$ means "for all." A statement such as "$\forall x$" followed by a condition on x means that all values of x satisfy the condition. The symbol $\exists$ means "there exists." A statement such as "$\exists x$" followed by a condition on x means that at least one value of x satisfies the condition. $\Box$

A.1.2 Sets

Definition A.2 A **set** is an unordered collection of elements. Sets will be denoted by symbols such as $\mathbb{S}$ and $\mathbb{P}$. We will use the symbols { and } to delimit the specification of the elements of a set.
The symbol $\in$ means "is an element of." Two sets are equal if every element of the first is an element of the second and every element of the second is an element of the first. The **set difference** $\mathbb{S} \setminus \mathbb{P}$ is the set of those elements of $\mathbb{S}$ that are not in $\mathbb{P}$.
The symbol $\subseteq$ means "is a non-strict subset of," which is to say that every element of the first set is also an element of the second set, but we allow for the possibility of equality of the two sets. If $\mathbb{P} \subseteq \mathbb{S}$ but $\mathbb{P} \neq \mathbb{S}$ then we say that $\mathbb{P}$ is a **strict subset** of $\mathbb{S}$. We use the symbol symbol $\subset$ to mean "is a strict subset of," which is to say that every element of the first set is also an element of the second set, but the sets are not equal. $\Box$

For example, $\mathbb{S} = \{1, 5, 2\}$ is the three element set consisting of the numbers 1, 2, 5. The statement "$1 \in \{1, 5, 2\}$" means that the number 1 is an element of the three element set $\{1, 5, 2\}$, (which is a true statement.) The statement "$\{1\} \subset \{1, 5, 2\}$" means that the set consisting of the number 1 is a strict subset of the three element set $\{1, 5, 2\}$, (which is a true statement.) Also, "$\{1, 5, 2\} \subseteq \{1, 5, 2\}$" means that the three element set $\{1, 5, 2\}$ is a non-strict subset of the three element set $\{1, 5, 2\}$, (which is also a true statement.) Finally, $\{1, 5, 2\} \setminus \{1, 3\} = \{5, 2\}$.

Definition A.3 We define $\mathbb{Z}$ to be the set of **integers**. The set $\mathbb{Z}_+$ is the set of non-negative integers, while $\mathbb{Z}_{++}$ is the set of strictly positive integers.
We define $\mathbb{R}$ to be the set of **real numbers**. The set $\mathbb{R}_+$ is the set of non-negative real numbers, while $\mathbb{R}_{++}$ is the set of strictly positive real numbers.
We define $\mathbb{K}$ to be the set of **complex numbers**. $\Box$

Definition A.4 Given two sets $\mathbb{S}$ and $\mathbb{P}$, the **Cartesian product** $\mathbb{S}\times\mathbb{P}$ is the set of all ordered pairs such that the first member of the pair is an element of $\mathbb{S}$ and the second member of the pair is an element of $\mathbb{P}$ [104, section 1.1]. We write $\mathbb{S}^n$ for the set of all ordered lists consisting of n (possibly non-distinct) elements of $\mathbb{S}$. We say the n-fold Cartesian product of $\mathbb{S}$ with itself. We write $\mathbb{S}^{m\times n}$ for the set of all ordered lists of m elements, each element itself being a member of $\mathbb{S}^n$. We say the $m \times n$ Cartesian product of $\mathbb{S}$ with itself. Given a collection of sets $\mathbb{S}_1, \dots, \mathbb{S}_n$, the Cartesian product of them is the set of all ordered lists, with each list consisting of elements of $\mathbb{S}_k$, $k = 1, \dots, n$, respectively, and is written $\prod_{k=1}^{n} \mathbb{S}_k$. $\Box$

Definition A.5 The n-fold Cartesian product of $\mathbb{R}_+$ with itself is called the **non-negative orthant** and is denoted $\mathbb{R}^n_+$. The n-fold Cartesian product of $\mathbb{R}_{++}$ with itself is called the **strictly positive orthant** and is denoted $\mathbb{R}^n_{++}$. If $\mathbb{M}$ is a set then $\mathbb{R}^{\mathbb{M}}$ is the set of all vectors having entries indexed by the elements in the set $\mathbb{M}$. $\Box$

A.1.3 Matrices, vectors, and scalars

Definition A.6 A **vector** is an element of a Cartesian product of sets, typically a Cartesian product of the form $\mathbb{R}^n$. We will usually think of the list that specifies the vector as being arranged as a **column** of n **entries** or **components**. We sometimes say a **column vector** to emphasize this. We can also define a **row vector** to be a list arranged as a row of entries.
A **matrix** is an element of an $m \times n$ Cartesian product and we can think of the list that specifies the matrix as being arranged as:

- m rows of n entries each, or
- n columns of m entries each.

A particular entry of a matrix or vector will be indicated by one or more subscripts on the symbol for the matrix or vector. By default, the subscripts are numbered consecutively from 1. For example $x \in \mathbb{R}^n$ will usually denote the vector consisting of the entries $x_1, \dots, x_n$. However, we will occasionally depart from this convention if it is more convenient to use non-consecutive numbering or to use other ways to list the entries of the vector. We will usually represent the entries of the vector by enclosing them with square brackets, so that in our example, $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$. Occasionally, we will represent a vector having two or three entries by enclosing them with round brackets and separating entries by commas. For example, (x, y, z) is a vector with entries x, y, and z.

For a matrix, the first subscript indexes the rows while the second subscript indexes the columns. Sometimes we will separate the first and second subscript with a comma to avoid ambiguity. The **transpose operator**, denoted by a superscript $\dagger$, interchanges rows and columns of a vector or matrix. $\square$

For example, $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ is the set of all ordered pairs of real numbers. Each element of $\mathbb{R}^2$ is a 2-vector. The entries of $x \in \mathbb{R}^2$ are x_1 and x_2, so that $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

The set $\mathbb{R}^n$ is the set of all ordered lists of n real numbers. Each element of $\mathbb{R}^n$ is an n-vector. The set $\mathbb{R}^n$ is also called n-dimensional Euclidean space, since it generalizes our notion of three-dimensional space for which "Euclidean geometry" applies. Moreover, for $x \in \mathbb{R}^n$, $x^\dagger \in \mathbb{R}^{1\times n}$ is the **transpose** of x; that is, $x^\dagger$ is a row vector with k-th entry equal to the k-th entry of the column vector x.

The set $\mathbb{R}^{m\times n}$ is the set of all $m \times n$ matrices of real numbers. Each element of this set is a matrix. For example, $\mathbb{R}^{2\times 3}$ is the set of all 2×3 matrices. For $A \in \mathbb{R}^{2\times 3}$, the entries of A are indexed as follows:

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}.$$

For $A \in \mathbb{R}^{m\times n}$, $A^\dagger$ is the **transpose** of A; that is, $A^\dagger$ is an $n \times m$ matrix with ℓk-th entry equal to the $k\ell$-th entry of A.

Definition A.7 A matrix $A \in \mathbb{R}^{m\times n}$ is **square** if it has the same number of rows and columns; that is, if $m = n$. $\square$

Definition A.8 The **diagonal** of a matrix $A \in \mathbb{R}^{m\times n}$ is the collection of entries A_{kk}, $k = 1, \ldots, \min\{m, n\}$, where $\min\{m, n\}$ means the smaller of m and n. A **diagonal matrix** is a matrix with:

- the same number of rows and columns (that is, a square matrix), and
- zero entries everywhere except on its diagonal.

$\square$

Definition A.9 A matrix $A \in \mathbb{R}^{m\times n}$ is **diagonally dominant** if:

$$\begin{aligned} \forall k, A_{kk} &\geq \sum_{\ell \neq k} |A_{k\ell}|, \\ \forall k, A_{kk} &\geq \sum_{\ell \neq k} |A_{\ell k}|. \end{aligned}$$

A matrix $A \in \mathbb{R}^{m \times n}$ is **strictly diagonally dominant** if:

$$\forall k, A_{kk} > \sum_{\ell \neq k} |A_{k\ell}|,$$
$$\forall k, A_{kk} > \sum_{\ell \neq k} |A_{\ell k}|.$$

□

Definition A.10 Let $A \in \mathbb{R}^{n \times n}$ be a square matrix. We define the **determinant** of A, denoted $\det(A)$, as follows. If $n = 1$ then $\det(A)$ is the single entry in the matrix itself. The determinant of an $n \times n$ matrix A can be calculated as the sum of n terms. The k-th term in the sum is given by the product of:

- $(-1)^{k+1}$,
- A_{1k}, and
- the determinant of the $(n-1) \times (n-1)$ sub-matrix of A obtained by deleting the first row and the k-th column of A.

□

The definition of determinant leads to a recursive algorithm for calculating the determinant having computational effort that increases with the factorial of n, that is with $n(n-1)(n-2)\cdots 1$, which we denote $n!$

We define some particular constant matrices and vectors in the following.

Definition A.11 The $n \times n$ **identity matrix** $\mathbf{I} \in \mathbb{R}^{n \times n}$ is a diagonal matrix with ones on the diagonal. We define $\mathbf{I}_k \in \mathbb{R}^n$ to be the k-th column of the identity matrix; that is, the vector with zeros everywhere except in the k-th entry, which has value 1.
We define $\mathbf{0}$ and $\mathbf{1}$, respectively, to be matrices or vectors of all zeros and all ones, respectively. The dimensions of $\mathbf{0}$ and $\mathbf{1}$ depend on the context. They will often be n-vectors of all zeros and all ones, respectively. □

A.1.4 Functions

Definition A.12 By $f : \mathbb{S} \to \mathbb{P}$ we mean that f is a **function** that takes elements from the **domain** set $\mathbb{S}$ and returns elements (function values) from the **range** set $\mathbb{P}$. That is, for each element $x \in \mathbb{S}$ there is a well-defined value $f(x) \in \mathbb{P}$. Sometimes we write $f(\bullet)$ for f to emphasize that f is a function. To define a function we must specify the value of the function for each element of its domain. □

In this book, we will always write $f(x)$ for the value of the function f at x and we will write f or $f(\bullet)$ for the function itself. That is, the symbol $f(x)$ *is not* a function: it is the value of the function f, evaluated at x. Usually, we think of the sets $\mathbb{S}$ and $\mathbb{P}$ as being disjoint; however, sometimes we may have $\mathbb{S} = \mathbb{P}$, $\mathbb{S} \subseteq \mathbb{P}$, or $\mathbb{P} \subseteq \mathbb{S}$, or sometimes one of the sets may be a subspace of the other. (See Definition A.51.)

Definition A.13 Let $f : \overline{\mathbb{S}} \to \mathbb{P}$ and suppose that $\mathbb{S} \subseteq \overline{\mathbb{S}}$. Then the **restriction** of f to $\mathbb{S}$ is the function that is defined on $\mathbb{S}$ and which matches f on this domain. We usually use the same symbol for a function and its restriction and distinguish the two by context. □

A.1.5 Alphabetical conventions

We will usually use Greek capital letters and italic Roman capital letters for matrices (and matrix-valued functions) and usually use Greek lower case and italic Roman lower case letters for vectors (and vector-valued functions.) We will use both capital and lower case letters for scalars (and scalar-valued functions.) The context will make clear whether a symbol stands for a scalar or stands for a vector or matrix. If we define a vector, $x \in \mathbb{R}^n$ say, then we will occasionally define the corresponding capital letter, X in this case, to be the **diagonal** matrix in $\mathbb{R}^{n\times n}$ with diagonal entries equal to the corresponding entries of x. That is, $X = \text{diag}\{x_\ell\}$. For example, if $x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \in \mathbb{R}^3$ then $X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \in \mathbb{R}^{3\times3}$. (The MATLAB function `diag` creates such a diagonal matrix from a vector.)

We will typically usually use Greek and italic Roman letters such as α, B, Γ that are near to the beginning of the Greek or Roman alphabets for **constants** and **parameters**; that is, scalars, vectors, or matrices that have entries that do not change or are held constant temporarily. We will use italic letters such as f, g, h that are further in to the Roman alphabet (and sometimes use their Greek cognates such as ϕ, γ, and η) for **functions**. We will occasionally not follow this convention. For example, we will occasionally use P and Q to stand for vectors, use β, J, and K to stand for functions, and use γ, η, ρ, and χ to stand for parameters and vectors.

We will use italic Roman letters such as j, k, ℓ, N and the Greek letter ν for counters. We will use k, ℓ, and, occasionally, j and i, to index entries of vectors. (We will usually, but not always, avoid indexing entries of vectors with the symbol i to avoid confusion with the symbol for electrical current. In the discussion of complex numbers, we write $\sqrt{-1}$ instead of i or j so that we can use the symbols i and j as counters.) The letters n, m, r, s will be reserved for the number of entries in particular vectors.

We will typically use italic letters such as x, y, z that are near to the end of the Roman alphabet and their Greek cognates for variables. The symbol Δ ligatured before a symbol for a variable will be used to denote a new variable that represents a *change* in the value of the original variable. For example, Δx denotes a change in x. The letters t, T, θ and calligraphic letters such as $\mathcal{A}$, $\mathcal{B}$, $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$ will be used in a variety of roles. An overline over or underline under a symbol for a variable means a constant of the same dimension that represents a bound on the variable or function represented by the symbol. For example, for $x \in \mathbb{R}^n$, the symbols $\underline{x} \in \mathbb{R}^n$ and $\overline{x} \in \mathbb{R}^n$ represent constant vectors that are lower and upper bounds for x.

In some case studies, we need to distinguish sub-systems or components. We will use arabic numerals and non-italic lower case letters such as a, b,. . ., f, g to distinguish these components. These symbols should not be confused with the corresponding italic symbols used for functions and vectors.

A.1.6 Superscripts and accents

We use superscripts and accents in several ways as specified in the following.

Definition A.14 To denote an **optimal** or **desired** value of a decision vector satisfying some criterion, we will use a superscript $\star$ (5-pointed star). For example, $x^\star$ will denote an optimal value of the vector $x \in \mathbb{R}^n$. We will occasionally consider the sensitivity of an optimal value with respect to the parameter χ. In these cases, we will abuse notation slightly and re-interpret, $x^\star$ say, to be a function representing the minimizer of a problem as a function of χ. We will use these conventions and natural generalizations of them throughout the book without further comment. $\square$

Definition A.15 We will use superscript $*$ (asterisk) to represent **complex conjugate**. $\square$

For definitions and theorems, we will often need to refer to one or more *typical* vectors or matrices. To distinguish the vectors and matrices, we will use superscripts and accents. For example:

- x, x', and x'' are three different vectors,
- x, $\tilde{x}$, and $\hat{x}$, are three different vectors, and
- if $\epsilon \in \mathbb{R}$ then we might distinguish a vector x^ϵ for each possible value of ϵ.

If $f : \mathbb{R}^n \to \mathbb{R}$ then we might distinguish a particular value or bound on the range of f by adding a superscript or accent. For example, we will usually write $f^\star$ for the optimal value of a function, where optimal is defined according to some criterion. As noted above, the individual **components** or **entries** of vectors are denoted by subscripts, so that x_k and x'_k are the k-th components of the vectors x and x', respectively.

Definition A.16 Let $x, x' \in \mathbb{R}^n$. We define the **vector relations** $=, \geq, >, <$, and $\leq$, respectively, by:

$$
\begin{aligned}
(x = x') &\quad\Leftrightarrow\quad (x_k = x'_k, \forall k = 1, \dots, n),\\
(x \geq x') &\quad\Leftrightarrow\quad (x_k \geq x'_k, \forall k = 1, \dots, n),\\
(x > x') &\quad\Leftrightarrow\quad (x_k > x'_k, \forall k = 1, \dots, n),\\
(x < x') &\quad\Leftrightarrow\quad (x_k < x'_k, \forall k = 1, \dots, n),\\
(x \leq x') &\quad\Leftrightarrow\quad (x_k \leq x'_k, \forall k = 1, \dots, n).
\end{aligned}
$$

That is, when a relation is used between vectors, the relation applies component-wise. □

Definition A.17 The set of **extended real numbers** is the set $\mathbb{R} \cup \{-\infty, \infty\}$ [104, section 2.3]. We *define* $-\infty$ and ∞ to have the following properties:

$$\begin{aligned} \forall \alpha \in \mathbb{R}, \quad & -\infty < \alpha < \infty, \\ & \alpha + \infty = \infty, \\ & \alpha + (-\infty) = -\infty. \end{aligned}$$

An **extended real function** f on $\mathbb{S} \subseteq \mathbb{R}^n$ is a function that, for each $x \in \mathbb{S}$, either takes on a value in $\mathbb{R}$ or takes on one of the special values $-\infty$ or ∞ [104, section 2.3]. That is, for each $x \in \mathbb{S}$, the value of $f(x)$ is an extended real number. We write that $f : \mathbb{S} \to \mathbb{R} \cup \{-\infty, \infty\}$. □

We will be careful never to *subtract* ∞ from ∞, nor to *add* ∞ to $-\infty$: these operations are not defined.

Definition A.18 We will use superscripts in parentheses to distinguish successive elements of a **sequence**. Usually, the sequences we consider will be the iterates produced by an iterative algorithm. The **initial guess** for an iterative algorithm will be denoted with a superscript (0), such as $x^{(0)}$; subsequent iterates will appear as $x^{(1)}, x^{(2)}, \ldots, x^{(\nu)}, \ldots$. To represent the set of all iterates, $\{x^{(0)}, x^{(1)}, \ldots\}$, that is, the complete sequence, we will write $\{x^{(\nu)}\}_{\nu=0}^{\infty}$.

We use the superscript parentheses to avoid confusion with **exponentiation**. If we want to represent the square of x_k, for example, we will write $(x_k)^2$, to clearly distinguish it from $x_k^{(2)}$, which is the value of the k-th component of the second iterate of the sequence $\{x^{(\nu)}\}_{\nu=0}^{\infty}$. Naturally, $\left(x_k^{(2)}\right)^3$ is the cube of the k-th component of the second iterate of this sequence.

Occasionally, we will need to consider an infinite sub-collection of elements of a sequence. For example, we might consider the sub-collection consisting of all the elements with even numbered iteration count: $\{x^{(0)}, x^{(2)}, x^{(4)}, \ldots, \}$. This is called a **sub-sequence** of the sequence $\{x^{(\nu)}\}_{\nu=0}^{\infty}$.

We will sometimes use superscript in parenthesis to distinguish elements of a finite collection.

□

A.2 Types of functions

We will classify functions by their functional form and by their properties. First, we will consider linear, affine, and quadratic functional forms and then we will consider polynomials and other functions.

A.2.1 Linear, affine, and quadratic

Definition A.19 A function $g : \mathbb{R}^n \to \mathbb{R}^m$ is **linear** if it is of the form:

$$\forall x \in \mathbb{R}^n, g(x) = Ax,$$

for some fixed $A \in \mathbb{R}^{m \times n}$. A function $g : \mathbb{R}^n \to \mathbb{R}^m$ is **affine** if it is of the form:

$$\forall x \in \mathbb{R}^n, g(x) = Ax - b,$$

for some fixed $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Recall that the ℓ-th entry of Ax is $\sum_{k=1}^n A_{\ell k} x_k$. In other words, the ℓ-th entry of Ax is determined by the ℓ-th row of A and by x; namely, it is the sum of the products of:

- the entries in the ℓ-th row of A, and
- the corresponding entries in x.

Then: $g_\ell(x) = \sum_{k=1}^n A_{\ell k} x_k - b_\ell$. $\Box$

Sometimes, authors use the word linear to refer both to linear and to affine functions.

Definition A.20 A function $f : \mathbb{R}^n \to \mathbb{R}$ is **quadratic** if it is of the form:

$$\begin{aligned} \forall x \in \mathbb{R}^n,\ f(x) &= \frac{1}{2} x^\dagger Q x + c^\dagger x + d, \\ &= \frac{1}{2} \sum_{k=1}^n \sum_{\ell=1}^n x_k Q_{k\ell} x_\ell + \sum_{k=1}^n c_k x_k + d, \end{aligned} \tag{A.1}$$

where:

- $Q \in \mathbb{R}^{n \times n}$,
- $c \in \mathbb{R}^n$, and
- $d \in \mathbb{R}$.

$\Box$

The factor $\frac{1}{2}$ in (A.1) is to simplify the functional form of the first derivative of the quadratic function. (See Section A.4.3.1 for definition of the first derivative.) If $Q = \mathbf{0}$ then the function is linear or affine. We often have that $d = 0$.

Definition A.21 A matrix $Q \in \mathbb{R}^{n \times n}$ is **symmetric** if $\forall k, \ell$, $Q_{k\ell} = Q_{\ell k}$. $\Box$

We can assume that Q in (A.1) is symmetric because, if it is not, we can replace it by $Q^\flat = \frac{1}{2}(Q + Q^\dagger)$, which is symmetric and yields the same value for the function, as Exercise A.1 shows.

A.2.2 Polynomial

Definition A.22 Let:

- $D \in \mathbb{Z}_+$ ($\mathbb{Z}_+$ is the set of non-negative integers; see Definition A.3), and
- $a_0, a_1, \dots, a_D \in \mathbb{R}$,

and define the function $g : \mathbb{R} \to \mathbb{R}$ by:

$$\forall x, g(x) = \sum_{k=0}^{D} a_k (x)^k.$$

This function is a **polynomial** of degree D in the single variable x. A polynomial is said to be affine, quadratic, cubic, or quartic if $D = 1, 2, 3$, or 4, respectively. $\square$

Linear, affine, quadratic, cubic, and quartic functions of a single variable are special classes of polynomials.

A.2.3 Other special functions

Definition A.23 A function $f : \mathbb{R}^n \to \mathbb{R}$ is **additively separable** if it is of the form:

$$\forall x \in \mathbb{R}^n,\ f(x) = \sum_{k=1}^{n} f_k(x_k),$$

where $f_k : \mathbb{R} \to \mathbb{R}, k = 1, \dots, n$. The function is **multiplicatively separable** if it is of the form:

$$\forall x \in \mathbb{R}^n,\ f(x) = \prod_{k=1}^{n} f_k(x_k).$$

$\square$

That is, a function is additively separable if it can be expressed as the sum of functions that each depend only on one entry of x. A function is multiplicatively separable if it can be expressed as the product of functions that each depend only on one entry of x. There are various other notions of separability. For example, a function is partially separable if it can be expressed as the sum of functions that each depend only on a particular sub-vector of x.

Definition A.24 A function $\eta^{\nearrow} : \mathbb{R} \to \mathbb{R}$ is **monotonically increasing** or **monotonically non-decreasing** if:

$$\forall x, x' \in \mathbb{R}, (x < x') \Rightarrow (\eta^{\nearrow}(x) \leq \eta^{\nearrow}(x')).$$

It is **strictly monotonically increasing** if:

$$\forall x, x' \in \mathbb{R}, (x < x') \Rightarrow (\eta^{\nearrow}(x) < \eta^{\nearrow}(x')).$$

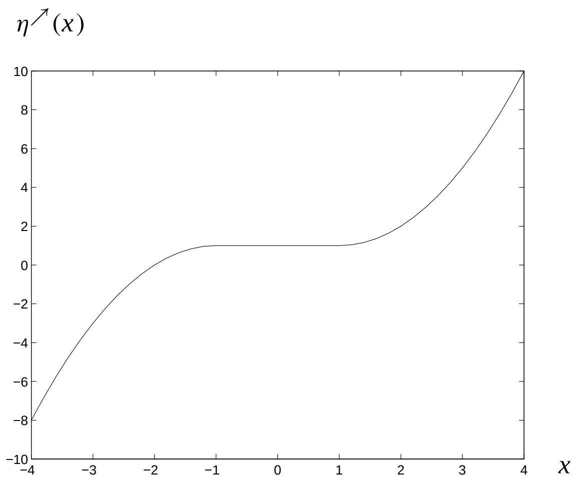

Fig. A.1. An example of a monotonically increasing function.

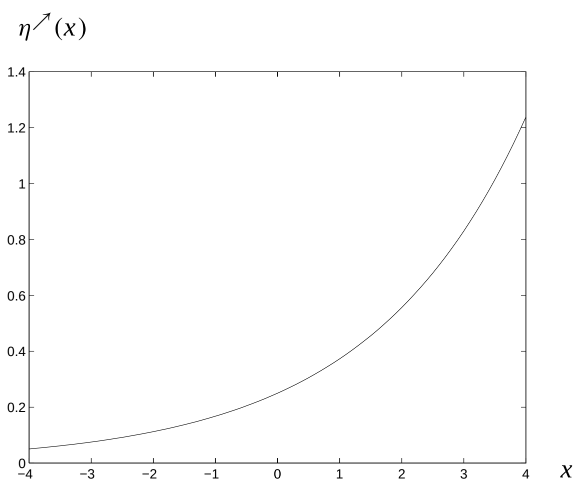

Fig. A.2. An example of a strictly monotonically increasing function.

Similarly, a function $\eta^{\searrow} : \mathbb{R} \to \mathbb{R}$ is **monotonically decreasing** or **monotonically non-increasing** if:

$$\forall x, x' \in \mathbb{R}, (x < x') \Rightarrow (\eta^{\searrow}(x) \geq \eta^{\searrow}(x')).$$

It is **strictly monotonically decreasing** if:

$$\forall x, x' \in \mathbb{R}, (x < x') \Rightarrow (\eta^{\searrow}(x) > \eta^{\searrow}(x')).$$

$\Box$

The superscripts $\nearrow$ and $\searrow$ are meant to graphically indicate the nature of monotonic functions. We can refer to $\eta^{\nearrow}$ as "eta-up" and refer to $\eta^{\searrow}$ as "eta-down" as mnemonics for their properties. Figures A.1 and A.2 show a monotonically increasing and a strictly monotonically increasing function, respectively.

Definition A.25 Let $\mathbb{S} \subseteq \overline{\mathbb{S}} \subseteq \mathbb{R}^n$, $\mathbb{P} \subseteq \mathbb{R}^r$, and $\tau : \mathbb{P} \to \overline{\mathbb{S}}$. (Recall that $\tau : \mathbb{P} \to \overline{\mathbb{S}}$ means that $\tau(\xi)$ is defined for each $\xi \in \mathbb{P}$ and that $\forall \xi \in \mathbb{P}, \tau(\xi) \in \overline{\mathbb{S}}$.) We say that τ is **onto** $\mathbb{S}$ if:

$$\forall x \in \mathbb{S}, \exists \xi \in \mathbb{P} \text{ such that } x = \tau(\xi).$$

We say that τ is **one-to-one** (or 1–1) if:

$$\forall \xi, \xi' \in \mathbb{P}, (\xi \neq \xi') \Rightarrow (\tau(\xi) \neq \tau(\xi')).$$

□

Definition A.26 We say that there is a **1–1 and onto correspondence** between two sets $\mathbb{P}$ and $\mathbb{S}$ if:

$$\exists \tau : \mathbb{P} \to \mathbb{S} \text{ such that } \tau \text{ is 1–1 and onto.}$$

□

Definition A.27 If $\tau : \mathbb{P} \to \mathbb{S}$ is 1–1 and onto then the **inverse** $\tau^{-1} : \mathbb{S} \to \mathbb{P}$ is defined by:

$$\forall x \in \mathbb{S}, \tau^{-1}(x) \text{ is the unique element } \xi \in \mathbb{P} \text{ such that } \tau(\xi) = x.$$

□

If $\tau : \mathbb{P} \to \mathbb{S}$ is 1–1 and onto then its inverse $\tau^{-1} : \mathbb{S} \to \mathbb{P}$ is also 1–1 and onto.

A.3 Norms

We define a measure of the length of a vector that generalizes our notion of length in space. This measure is called a **norm** [104, section 10.1]. We then define the notion of the norm of a matrix.

A.3.1 Vector

Definition A.28 A **norm** (or **vector norm**) on $\mathbb{R}^n$ is a function, $\|\bullet\| : \mathbb{R}^n \to \mathbb{R}$, with the following properties:

(i) $\forall x \in \mathbb{R}^n, \|x\| \geq 0$,
(ii) $\forall x \in \mathbb{R}^n, (\|x\| = 0) \Leftrightarrow (x = \mathbf{0})$,
(iii) $\forall x, y \in \mathbb{R}^n, \|x + y\| \leq \|x\| + \|y\|$,
(iv) $\forall x \in \mathbb{R}^n, \forall \alpha \in \mathbb{R}, \|\alpha x\| = |\alpha| \, \|x\|$.

□

The most familiar example of a norm on $\mathbb{R}^n$ is the **Euclidean length**, usually denoted $\|\bullet\|_2$ and defined by:

$$\forall x \in \mathbb{R}^n, \|x\|_2 = \sqrt{\sum_{k=1}^{n} (x_k)^2}.$$

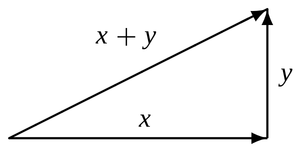

Fig. A.3. Illustration of the triangle inequality in two dimensions.

This norm is also called the L_2 norm and is the same as our intuitive notion of length in 1, 2, or 3 dimensions. Property (iii) of a norm is called the **triangle inequality** because it says that the sum of the lengths of two sides of a triangle exceeds the length of the other side. The triangle inequality is illustrated for $n = 2$ in Figure A.3. In this figure, the sum of the lengths of the vertical and horizontal sides of the triangle exceeds the length of the oblique side. The same observation applies for each of the other two pairs of sides. Properties (iii) and (iv) imply that a norm is a continuous function. (See Definition A.35 and Exercise A.8.)

There are many other norms, such as:

- the L_1 **norm** $\|\bullet\|_1$ defined by:

$$\forall x \in \mathbb{R}^n, \|x\|_1 = \sum_{k=1}^{n} |x_k|,$$

- the L_∞ or **infinity norm** $\|\bullet\|_\infty$ defined by:

$$\forall x \in \mathbb{R}^n, \|x\|_\infty = \max_{k=1,\dots,n} \{|x_k|\},$$

  and
- **weighted norms** $\|\bullet\|_W$ defined in terms of a non-singular weighting matrix (see Definition A.49) $W \in \mathbb{R}^{n \times n}$ and any other norm $\|\bullet\|$ on $\mathbb{R}^n$ by:

$$\forall x \in \mathbb{R}^n, \|x\|_W = \|Wx\|.$$

The choice of norm depends on the application. However, for any norms $\|\bullet\|$ and $\|\bullet\|'$ on $\mathbb{R}^n$, there are constants $\underline{\kappa}, \overline{\kappa} \in \mathbb{R}_{++}$ such that:

$$\forall x \in \mathbb{R}^n, \underline{\kappa}\,\|x\| \leq \|x\|' \leq \overline{\kappa}\,\|x\|.$$

(See Exercise A.2.) In more general spaces than $\mathbb{R}^n$ this is not necessarily true.

In several theorems, our results will be stated in terms of norms. Usually, the result is independent of the particular choice of norm. In this case we will use the symbol $\|\bullet\|$ to denote any particular norm. Of course, we must use the same norm consistently throughout the theorem. Occasionally we will use $\|\bullet\|$ to refer to norms in two different spaces, say $\mathbb{R}^n$ and $\mathbb{R}^m$. This is a slight abuse of notation, since the norms are, strictly speaking, different and should be distinguished notationally. Naturally, we must, for example, use the norm consistently for $\mathbb{R}^n$ and consistently for $\mathbb{R}^m$. However, unless otherwise specified, the norm on $\mathbb{R}^n$ could be, say, $\|\bullet\|_1$, while the norm on $\mathbb{R}^m$ could be, say, $\|\bullet\|_2$.

A.3.2 Matrix

We would also like to "measure" matrices. We make the following definition.

Definition A.29 A **norm** (or **matrix norm**) on $\mathbb{R}^{m\times n}$ is a function, $\|\bullet\| : \mathbb{R}^{m\times n} \to \mathbb{R}$, with the following properties:

(i) $\forall A \in \mathbb{R}^{m\times n}, \|A\| \geq 0$,
(ii) $\forall A \in \mathbb{R}^{m\times n}, (\|A\| = 0) \Leftrightarrow (A = \mathbf{0})$,
(iii) $\forall A, B \in \mathbb{R}^{m\times n}, \|A + B\| \leq \|A\| + \|B\|$,
(iv) $\forall A \in \mathbb{R}^{m\times n}, \forall \alpha \in \mathbb{R}, \|\alpha A\| = |\alpha| \, \|A\|$.

□

We often use the particular matrix norm described in the following.

Definition A.30 Suppose we have two vector norms $\|\bullet\|$ defined on $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, and a matrix $A \in \mathbb{R}^{m\times n}$. Then the **induced matrix norm** $\|\bullet\| : \mathbb{R}^{m\times n} \to \mathbb{R}$ is defined by:

$$\forall A \in \mathbb{R}^{m\times n}, \|A\| = \max_{\|x\|=1} \|Ax\| , \tag{A.2}$$

where:

- the norm in $\|x\|$ is the norm on $\mathbb{R}^n$,
- the norm in $\|Ax\|$ is the norm on $\mathbb{R}^m$, and
- the norm in $\|A\|$ is the induced matrix norm that is being defined.

□

(The maximum on the right-hand side of (A.2) exists by Theorem 2.1 since the max is over a bounded set (see Definition A.46) and the norm is a continuous function. See Definition A.35 and Exercise A.8.) An induced matrix norm is a matrix norm according to Definition A.29. (See Exercise A.4.) If the norms on $\mathbb{R}^n$ and $\mathbb{R}^m$ are, say, both the L_2 norms or both the L_1 norms, then we will typically use the same symbol for the norm on $\mathbb{R}^n$, the norm on $\mathbb{R}^m$, and the induced matrix norm. The appropriate norm will be clear from the context. However, if the norms on $\mathbb{R}^n$ and $\mathbb{R}^m$ are different then the symbols should be more carefully distinguished.

We have the following.

Lemma A.1 *Suppose that we have three vector norms $\|\bullet\|$ defined on $\mathbb{R}^n, \mathbb{R}^m$, and $\mathbb{R}^r$, respectively. Then:*

$$\forall A \in \mathbb{R}^{m\times n}, x \in \mathbb{R}^n, \|Ax\| \leq \|A\| \, \|x\| , \tag{A.3}$$
$$\forall A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times r}, \|AB\| \leq \|A\| \, \|B\| ,$$

where each matrix norm is induced by the corresponding pair of vector norms.

Proof First observe that:

$$\begin{aligned}
\forall x \neq \mathbf{0}, \left\| \frac{1}{\|x\|} x \right\| &= \frac{1}{\|x\|} \|x\|, \text{ by Property (iv) of norms, since } |1/\|x\|| = 1/\|x\|, \\
&= 1.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\|Ax\| &= \left\| \|x\| A \frac{1}{\|x\|} x \right\|, \text{ multiplying and dividing by a constant,} \\
&= \|x\| \left\| A \frac{1}{\|x\|} x \right\|, \text{ by Property (iv) of norms, since } |\ \|x\|\ | = \|x\|, \\
&\leq \|x\| \|A\|, \text{ by definition of } \|A\|, \text{ since } \left\| \tfrac{1}{\|x\|} x \right\| = 1.
\end{aligned}$$

Now let $\|y\| = 1$. Then,

$$\begin{aligned}
\|ABy\| &\leq \|A\| \|By\|, \text{ by (A.3) applied to } A \in \mathbb{R}^{m \times n} \text{ and } By \in \mathbb{R}^n, \\
&\leq \|A\| \|B\| \|y\|, \text{ by (A.3) applied to } B \in \mathbb{R}^{m \times r} \text{ and } y \in \mathbb{R}^r, \\
&= \|A\| \|B\|, \text{ since } \|y\| = 1.
\end{aligned}$$

Taking the maximum of the left-hand side over all vectors having norm 1, we obtain from (A.2) that $\|AB\| \leq \|A\| \|B\|$. □

If the norms on $\mathbb{R}^n$ and $\mathbb{R}^m$ are both L_2 norms, then we write $\|\bullet\|_2$ for the induced matrix norm and call it the L_2 **matrix norm**. For any $A \in \mathbb{R}^{m \times n}$, $\|A\|_2$ is equal to the maximum **singular value** of A [45, section 2.2.5.5][55, appendix]. The singular values of A are the non-negative square roots of the eigenvalues of $A^{\dagger}A$. If $A \in \mathbb{R}^{n \times n}$ is symmetric then $\|A\|_2$ is equal to the largest of the absolute values of the eigenvalues of A [45, section 2.2.5.5][55, appendix]. Recall the following definition.

Definition A.31 Let $A \in \mathbb{R}^{n \times n}$ be square and suppose that we can find $\lambda \in \mathbb{K}$ and $\xi \in \mathbb{K}^n$ such that $A\xi = \lambda\xi$. Then λ is called an **eigenvalue** and ξ is called an **eigenvector** of A. □

In general, there are n eigenvalues for an $n \times n$ matrix, given by the solution of the **characteristic equation**:

$$\det(A - \mathbf{I}\lambda) = 0.$$

Definition A.32 Vector norms $\|\bullet\|$ on $\mathbb{R}^n$ and $\mathbb{R}^m$ and a matrix norm $\|\bullet\|'$ are called **compatible** if:

$$\forall x \in \mathbb{R}^n, \forall A \in \mathbb{R}^{m \times n}, \|Ax\| \leq \|A\|' \|x\|.$$

□

By definition, vector norms on $\mathbb{R}^n$ and $\mathbb{R}^m$ and the corresponding induced matrix norm are compatible. However, there are matrix norms which are not compatible with any vector norm. For example, the **Frobenius norm**:

$$\|A\|_F = \left(\sum_{k=1}^{m}\sum_{\ell=1}^{n}(A_{k\ell})^2\right)^{\frac{1}{2}}$$

is not compatible with any vector norms. More details on matrix norms are contained in [45, section 2.2.4.2].

A.4 Limits

We discuss limiting properties of sequences and of functions.

A.4.1 Convergence and limits

Sequences have limiting properties embodied in the following.

Definition A.33 Let $\|\bullet\|$ be a norm on $\mathbb{R}^n$. (See Definition A.28 for the definition of norm.) Let $\{x^{(\nu)}\}_{\nu=0}^{\infty}$ be a sequence of vectors in $\mathbb{R}^n$. Then, the sequence $\{x^{(\nu)}\}_{\nu=0}^{\infty}$ **converges** to a **limit** $x^\star$ if:

$$\forall \epsilon > 0, \exists N \in \mathbb{Z}_+ \text{ such that } (\nu \in \mathbb{Z}_+ \text{ and } \nu \geq N) \Rightarrow \left\|x^{(\nu)} - x^\star\right\| \leq \epsilon.$$

We write $\lim_{\nu\to\infty} x^{(\nu)} = x^\star$ or $\lim\limits_{\nu\to\infty} x^{(\nu)} = x^\star$ and call $x^\star$ the **limit** of the sequence $\{x^{(\nu)}\}_{\nu=0}^{\infty}$. □

Definition A.34 A sequence $\{x^{(\nu)}\}_{\nu=0}^{\infty}$ has an **accumulation point** $x^\star$ if some sub-sequence of the sequence converges to $x^\star$. □

A.4.2 Continuity

Definition A.35 A function $g : \mathbb{R}^n \to \mathbb{R}^m$ is **continuous at** $x^\star$ if there are any norms $\|\bullet\|$ on $\mathbb{R}^n$ and $\mathbb{R}^m$ such that:

$$\forall \epsilon > 0, \exists \delta > 0 \text{ such that } \left(\left\|x^\star - x\right\| \leq \delta\right) \Rightarrow \left(\left\|g(x^\star) - g(x)\right\| \leq \epsilon\right). \tag{A.4}$$

A function is **continuous on** $\mathbb{S} \subseteq \mathbb{R}^n$ if it is continuous at $x^\star$ for every $x^\star \in \mathbb{S}$. If a function is continuous on $\mathbb{S} = \mathbb{R}^n$, then it is said to be **continuous** or **continuous everywhere**. □

Notice that by Exercise A.2, Part (iv), for a given ϵ, the largest value of δ that satisfies (A.4) will depend on which norm is used on $\mathbb{R}^n$; however, it can be shown that the property of continuity of a function $g : \mathbb{R}^n \to \mathbb{R}^m$ is independent of the choice of norm on $\mathbb{R}^n$ and $\mathbb{R}^m$. In more general spaces than $\mathbb{R}^n$, this is not true [82, section 2-7].

A.4.3 Differentiation

A.4.3.1 First derivative

Definition A.36 We say that a function $f : \mathbb{R} \to \mathbb{R}$ is **differentiable** at $x^\star$ with respect to x or its **first derivative** with respect to x exists at $x^\star$ if the following limit exists:

$$\lim_{\delta \to 0} \frac{f(x^\star + \delta) - f(x^\star)}{\delta}.$$

The value of the limit is denoted $\frac{df}{dx}(x^\star)$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is **partially differentiable** at $x^\star$ if, for $k = 1, \ldots, n$, the first derivatives with respect to x_k all exist. We write $\frac{\partial f}{\partial x_k}(x^\star)$ for the first derivative with respect to x_k, $k = 1, \ldots, n$, and call them the **first partial derivatives** at $x^\star$. A function $g : \mathbb{R}^n \to \mathbb{R}^m$ is partially differentiable if each function g_ℓ, $\ell = 1, \ldots, m$, is partially differentiable.
Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ are partially differentiable at $x^\star$. That is, suppose that $\frac{\partial f}{\partial x_k}(x^\star)$ exists for each k and suppose that $\frac{\partial g_\ell}{\partial x_k}(x^\star)$ exists for each k and ℓ.

Then the **derivative** and **gradient** of f at $x^\star$, symbols $\frac{\partial f}{\partial x}(x^\star)$ and $\nabla f(x^\star)$, respectively, are defined as follows:

- $\frac{\partial f}{\partial x}(x^\star) \in \mathbb{R}^{1 \times n}$ is the *row* vector with k-th entry equal to $\frac{\partial f}{\partial x_k}(x^\star)$, and
- $\nabla f(x^\star) \in \mathbb{R}^n$ is the *column* vector with k-th entry equal to $\frac{\partial f}{\partial x_k}(x^\star)$.

We have that $\frac{\partial f}{\partial x}(x^\star) = [\nabla f(x^\star)]^\dagger$.

Furthermore, the **derivative** and **gradient** of g, symbols $\frac{\partial g}{\partial x}$ and ∇g, respectively, are defined as follows:

- $\frac{\partial g}{\partial x}(x^\star) \in \mathbb{R}^{m \times n}$ is the matrix with ℓk-th entry equal to $\frac{\partial g_\ell}{\partial x_k}(x^\star)$, and
- $\nabla g(x^\star) \in \mathbb{R}^{n \times m}$ is the matrix with $k\ell$-th entry equal to $\frac{\partial g_\ell}{\partial x_k}(x^\star)$.

That is, $\frac{\partial g}{\partial x}(x^\star) = [\nabla g(x^\star)]^\dagger$.
If the partial derivatives exist for all points in the set $\mathbb{S} \subseteq \mathbb{R}^n$ then the function is said to be partially differentiable on $\mathbb{S}$. We write $\frac{\partial f}{\partial x}$, ∇f, $\frac{\partial g}{\partial x}$, and ∇g for the functions whose values at each $x^\star \in \mathbb{S}$ is given by $\frac{\partial f}{\partial x}(x^\star)$, $\nabla f(x^\star)$, $\frac{\partial g}{\partial x}(x^\star)$, and $\nabla g(x^\star)$, respectively. If the partial derivatives exist for all points in $\mathbb{R}^n$ then the function is said to be partially differentiable or partially differentiable everywhere. (For the distinction between a partially differentiable function and a differentiable function, see [72, section 2.3].)

The symbol ∇ is sometimes pronounced "del." The matrix $\frac{\partial g}{\partial x}$ is also called the **Jacobian** and we often use the symbol J for the Jacobian of the function g.
$\square$

Definition A.37 If $f : \mathbb{R}^n \to \mathbb{R}$ is partially differentiable with continuous partial derivatives and $x^\star \in \mathbb{R}^n$, $\Delta x \in \mathbb{R}^n$, then the function $\phi : \mathbb{R} \to \mathbb{R}$ defined by:

$$\forall t \in \mathbb{R}, \phi(t) = f(x^\star + t\Delta x)$$

is a differentiable function. Moreover, by the chain rule [72, section 2.4]:

$$\frac{d\phi}{dt}(0) = \nabla f(x^\star)^\dagger \Delta x.$$

We call $\nabla f(x^\star)^\dagger \Delta x$ the **directional derivative** of f at $x^\star$ in the direction Δx since is evaluates the rate of change of f in the direction Δx from $x^\star$. $\square$

If f is partially differentiable at a point $x^\star \in \mathbb{R}^n$ but its partial derivatives are not continuous at $x^\star$ then the function ϕ in Definition A.37 may not be differentiable. (See Exercise A.9.)

A.4.3.2 Second derivative

Definition A.38 A **second derivative** is a derivative of a derivative function. For a function $f : \mathbb{R}^n \to \mathbb{R}$ we write $\frac{\partial^2 f}{\partial x_\ell \partial x_k}$ for the derivative with respect to x_ℓ of the derivative with respect to x_k of f. If these functions exist for each ℓ and k then we say that the function is twice partially differentiable. We then define $\frac{\partial^2 f}{\partial x^2} : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ by:

$$\forall \ell, k, \left[\frac{\partial^2 f}{\partial x^2}\right]_{\ell k} = \frac{\partial^2 f}{\partial x_\ell \partial x_k}.$$

We call $\frac{\partial^2 f}{\partial x^2}$ the **Hessian** of f and we also write $\nabla^2 f$ and $\nabla^2_{xx} f$ for $\frac{\partial^2 f}{\partial x^2}$. $\square$

The Hessian of a function is the same as the Jacobian of its gradient. Exercise A.10 shows the reason for the $\frac{1}{2}$ in Definition A.20 of a quadratic function. Exercise A.10 shows that if f is quadratic then its Hessian is constant. If f is approximately quadratic, then its Hessian is approximately constant.

A.4.3.3 Symbolic conventions

Symbols and conventions for functions and derivatives are often confusing. We will use the following convention. Each function we introduce will be defined in terms of a "dummy variable." The dummy variable is the argument as specified in the definition of the function. We must specify the value of the function for each possible value that the dummy variable can take on in the domain of the function;

that is, the dummy variable is running over all possible values. We will then avoid using the dummy variable in any role where it is thought of as a constant or a particular value.

For example, suppose that g were defined using the dummy variable x. When we refer to the function we will write either g or $g(\bullet)$, omitting the dummy variable. To indicate g evaluated at a particular point x' we write $g(x')$.

To indicate the derivative of the function evaluated at a point x' we write $\frac{\partial g}{\partial x}(x')$. The derivative function will be denoted $\frac{\partial g}{\partial x}$ or $\frac{\partial g}{\partial x}(\bullet)$. We will avoid using x to stand, at the same time, for the dummy variable *and* for a particular point in an expression because of the difficulty in distinguishing:

- the use of x as the dummy variable in $\frac{\partial g}{\partial x}$ from
- the use of x in the argument of $\frac{\partial g}{\partial x}(x)$.

To see this issue, consider the function $g : \mathbb{R} \to \mathbb{R}$ defined by:

$$\forall x, g(x) = (x)^3.$$

Then, $\frac{\partial g}{\partial x}$ is the function defined by:

$$\forall x, \frac{\partial g}{\partial x}(x) = 3(x)^2.$$

If we write:

$$\frac{\partial g}{\partial x}((x)^2), \tag{A.5}$$

then we mean the function $\frac{\partial g}{\partial x}$ evaluated at the point $(x)^2$, which is $3((x)^2)^2 = 3(x)^4$. However, we interpret the similar-looking expression $\frac{\partial}{\partial x}[g((x)^2)]$ as meaning:

$$\begin{aligned} \frac{\partial}{\partial x}[g((x)^2)] &= \frac{\partial g}{\partial x}((x)^2) \times \frac{\partial}{\partial x}[(x)^2], \\ &= 3(x)^4 \times 2x, \\ &= 6(x)^5, \end{aligned} \tag{A.6}$$

using the chain rule [72, section 2.4]. Because it is easy to confuse (A.5) and (A.6) we will usually try to avoid expressions like them and we will typically use the ∇ notation to denote the gradient function.

A.4.4 Integration

A.4.4.1 Fundamental theorem of integral calculus

Theorem A.2 *Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function and let $a, b \in \mathbb{R}$. Then:*

$$f(b) - f(a) = \int_{t=a}^{b} \frac{df}{dt}(t)\, dt.$$

Proof See [114, section 4-8]. □

A.4.4.2 Integration of non-negative function

Theorem A.3 *Let $f : \mathbb{R} \to \mathbb{R}_+$ be continuous and let $a, b \in \mathbb{R}$. Then:*

$$\int_{t=a}^{b} f(t)\, dt \geq 0.$$

If $f(t)$ is strictly positive for $a < t < b$ then the integral is strictly positive.

Proof See [114, section 4-8] and Exercise A.11. □

A.5 Sets

A.5.1 Notation

It is often convenient to define sets by collecting together all those elements from another set, such as $\mathbb{R}$ or $\mathbb{R}^n$, that have a particular property. We formalize this in the following.

Definition A.39 Let $\Theta : \mathbb{S} \to \{\text{true, false}\}$ be a function that evaluates to either true or false. By $\{x \in \mathbb{S} | \Theta(x)\}$ we mean the subset of $\mathbb{S}$ consisting of all those elements x such that $\Theta(x)$ is true. The function Θ is often expressed "loosely." □

For example, $\{x \in \mathbb{R}^2 | -1 \leq x_1 \leq 1\}$ means the set of all two-vectors such that the first entry of the two vector, namely x_1, has a value that lies between -1 and 1.

If the dummy variable in the definition of Θ and the set $\mathbb{S}$ are clear from context, then we sometimes omit the "$x \in \mathbb{S}|$." For example, if the context is clear, we might write $\{-1 \leq x \leq 1\}$ for $\{x \in \mathbb{R} | -1 \leq x \leq 1\}$. If there are multiple conditions in the definition of the set then these are separated by commas. They should be interpreted as meaning "and" or "intersection." For example, $\{x \in \mathbb{R}^n | g(x) = \mathbf{0}, h(x) \leq \mathbf{0}\}$ means the set of vectors x in $\mathbb{R}^n$ such that $g(x) = \mathbf{0}$ and $h(x) \leq \mathbf{0}$.

A.5.2 Open and closed

Definition A.40 A point $x^{\mathrm{c}} \in \mathbb{R}^n$ is called a **point of closure** or a **limit point** of a set $\mathbb{S} \subseteq \mathbb{R}^n$ if there is a norm $\|\bullet\|$ such that:

$$\forall \epsilon > 0, \exists x^{\epsilon} \in \mathbb{S} \text{ such that } \left\|x^{\epsilon} - x^{\mathrm{c}}\right\| \leq \epsilon.$$

A point $x^{\mathrm{i}} \in \mathbb{R}^n$ is called an **interior point** of a set $\mathbb{S} \subseteq \mathbb{R}^n$ if there is a norm $\|\bullet\|$ such that:

$$\exists \epsilon > 0 \text{ such that } \forall x \in \mathbb{R}^n, \left(\left\|x^{\mathrm{i}} - x\right\| \leq \epsilon\right) \Rightarrow (x \in \mathbb{S}).$$

The set of all limit points of $\mathbb{S}$ is denoted by $\mathrm{cl}(\mathbb{S})$. The set of all interior points of $\mathbb{S}$ is denoted by $\mathrm{int}(\mathbb{S})$ and is called its **interior**. $\square$

Any point in $\mathbb{S}$ is also a limit point of $\mathbb{S}$, but in general some limit points of $\mathbb{S}$ may not be contained in $\mathbb{S}$. That is, $\mathbb{S} \subseteq \mathrm{cl}(\mathbb{S})$. Any interior point of $\mathbb{S}$ is contained in $\mathbb{S}$, but in general some points of $\mathbb{S}$ are not interior points of $\mathbb{S}$. That is, $\mathrm{int}(\mathbb{S}) \subseteq \mathbb{S}$.

Definition A.41 A set $\mathbb{S} \subseteq \mathbb{R}^n$ is **closed** if it contains all its limit points. That is, $\mathbb{S}$ is closed if $\mathrm{cl}(\mathbb{S}) = \mathbb{S}$. A set $\mathbb{S} \subseteq \mathbb{R}^n$ is **open** if $(\mathbb{R}^n \setminus \mathbb{S})$ is closed or, equivalently, if every point in $\mathbb{S}$ is an interior point of $\mathbb{S}$. That is, $\mathbb{S}$ is open if $\mathrm{int}(\mathbb{S}) = \mathbb{S}$. $\square$

Definition A.42 The **boundary** of a set $\mathbb{S} \subseteq \mathbb{R}^n$ is defined to be the set $(\mathrm{cl}(\mathbb{S}) \setminus \mathrm{int}(\mathbb{S}))$. $\square$

For a point x^{b} on the boundary of $\mathbb{S} \subseteq \mathbb{R}^n$ there are points in $\mathbb{S}$ that are arbitrarily close to x^{b} and points not in $\mathbb{S}$ that are arbitrarily close to x^{b}. A closed set contains its boundary. For example, consider a "closed ball" as defined in the following.

Definition A.43 A **closed ball** of radius $\rho \in \mathbb{R}_{++}$ about a point $x^{(0)} \in \mathbb{R}^n$ is the set:

$$\left\{x \in \mathbb{R}^n \,\middle|\, \left\|x - x^{(0)}\right\| \leq \rho\right\}.$$

$\square$

A closed ball is (not surprisingly) a closed set and contains its boundary. By Definition A.40, for any interior point x^{i} of a set $\mathbb{S}$, we can find a closed ball of some radius $\epsilon > 0$ about x^{i} that is contained in $\mathbb{S}$. We can also define an "open ball."

Definition A.44 An **open ball** of radius $\rho \in \mathbb{R}_{++}$ about a point $x^{(0)} \in \mathbb{R}^n$ is the set:

$$\left\{x \in \mathbb{R}^n \,\middle|\, \left\|x - x^{(0)}\right\| < \rho\right\}.$$

$\square$

The interior of a closed ball is the corresponding open ball.

Definition A.45 An open set in $\mathbb{R}^n$ containing a point $x^{(0)}$ is called a **neighborhood** of $x^{(0)}$. $\square$

An example of a neighborhood of $x^{(0)}$ is an open ball of radius $\rho > 0$ about $x^{(0)}$.

Definition A.46 A set $\mathbb{S} \in \mathbb{R}^n$ is **bounded** if there exists $\rho \in \mathbb{R}_+$ and a norm $\|\bullet\|$ such that $\forall x \in \mathbb{S}$, $\|x\| \leq \rho$. □

A closed ball is bounded. An open ball is bounded.

A.5.3 Projections

Definition A.47 Let $\mathbb{S} \subseteq \mathbb{R}^n$, let $n' \leq n$, and let $\mathbb{P} \subseteq \mathbb{R}^{n'}$ be defined by:

$$\mathbb{P} = \left\{ \xi \in \mathbb{R}^{n'} \middle| \exists x \in \mathbb{S} \text{ such that } \xi_k = x_{k+n-n'}, k = 1, \ldots n' \right\}.$$

The set $\mathbb{P}$ is called the **projection of $\mathbb{S}$ onto the last n' components of $\mathbb{R}^n$**. If $n' = 1$ then we call $\mathbb{P}$ the projection of $\mathbb{S}$ on the last component of $\mathbb{R}^n$. Similarly, we can define the projection onto any other subset of the components. □

For example, if $\mathbb{S} \subseteq \mathbb{R}^2$ is the closed ball of radius 1 centered at $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ then the projection of $\mathbb{S}$ onto the last component of $\mathbb{R}^2$ is the set $\mathbb{P} = \{x_1 \in \mathbb{R} | -1 \leq x_1 \leq 1\} \subseteq \mathbb{R}$.

Definition A.48 Let $\|\bullet\|$ be a norm, $\mathbb{S} \subseteq \mathbb{R}^n$, and $\hat{x} \in \mathbb{R}^n$. Then the **projection of $\hat{x}$ on $\mathbb{S}$** is the set $\operatorname{argmin}_{x \in \mathbb{S}} \left\{ \left\| x - \hat{x} \right\| \right\}$ [15, sections 6.1 and 8.1]. □

A.6 Properties of matrices

A.6.1 Singular and non-singular matrices

A.6.1.1 Definitions

Definition A.49 A square matrix $A \in \mathbb{R}^{n \times n}$ is **invertible** if there exists another matrix in $\mathbb{R}^{n \times n}$, (which we write A^{-1} and call the **inverse**) that satisfies:

$$A^{-1} A = A A^{-1} = \mathbf{I}.$$

An invertible matrix is also referred to as **non-singular**. If no inverse exists, then A is called **singular**. □

Definition A.50 Let $A \in \mathbb{R}^{m \times n}$. Then we define the following.

- The **range space** of A is the set $\mathcal{R}(A) = \{y \in \mathbb{R}^m | \exists x \in \mathbb{R}^n \text{ such that } y = Ax\}$. (We often abbreviate this expression by writing: $\{Ax \in \mathbb{R}^m | x \in \mathbb{R}^n\}$, where it is understood that the set contains the values $y = Ax$ for all $x \in \mathbb{R}^n$.)
- The **null space** of A is the set $\mathcal{N}(A) = \{x \in \mathbb{R}^n | Ax = \mathbf{0}\}$.

□

For example, consider the following.

- If $A = \mathbf{0}$ then $\mathcal{R}(A) = \{Ax \in \mathbb{R}^m | x \in \mathbb{R}^n\} = \{\mathbf{0}\}$ and $\mathcal{N}(A) = \{x \in \mathbb{R}^n | Ax = \mathbf{0}\} = \mathbb{R}^n$.
- If $A = \mathbf{1}^\dagger$ then $\mathcal{R}(A) = \{Ax \in \mathbb{R}^m | x \in \mathbb{R}^n\} = \mathbb{R}$ and $\mathcal{N}(A) = \{x \in \mathbb{R}^n | Ax = \mathbf{0}\} = \{x \in \mathbb{R}^n | \sum_{k=1}^n x_k = 0\}$.
- If $A = \mathbf{I}$ then $\mathcal{R}(A) = \{Ax \in \mathbb{R}^m | x \in \mathbb{R}^n\} = \mathbb{R}^n$ and $\mathcal{N}(A) = \{x \in \mathbb{R}^n | Ax = \mathbf{0}\} = \{\mathbf{0}\}$.

Since $A\mathbf{0} = \mathbf{0}$ for any matrix A, the zero vector is an element of both the range space and the null space of any matrix.

For any $A \in \mathbb{R}^{m \times n}$, we have the somewhat surprising result that any vector in $\mathbb{R}^n$ can be expressed as the sum of ([55, section A.15]):

- an element of the range space of $A^\dagger$, plus
- an element of the null space of A.

That is, we have the following.

Theorem A.4 *Let $A \in \mathbb{R}^{m \times n}$. Then,*

$$\forall x \in \mathbb{R}^n, \exists \lambda \in \mathbb{R}^m, \exists z \in \mathbb{R}^n \textit{ with } Az = \mathbf{0} \textit{ such that } x = z + A^\dagger \lambda.$$

Proof See [55, section A.15] and Exercise 5.47. □

Definition A.51 A **vector subspace** of $\mathbb{R}^n$ is a set $\mathbb{S} \subseteq \mathbb{R}^n$ with the following properties:

(i) $\forall x, x' \in \mathbb{S}, x + x' \in \mathbb{S}$,
(ii) $\forall x \in \mathbb{S}, \forall \alpha \in \mathbb{R}, \alpha x \in \mathbb{S}$.

□

The set $\mathbb{R}^n$ is a vector subspace of itself. The null space $\mathcal{N}(A)$ and range space $\mathcal{R}(A)$ of a matrix $A \in \mathbb{R}^{m \times n}$ are vector subspaces of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively.

Definition A.52 Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. A set of the form $\{x \in \mathbb{R}^n | Ax = b\}$ is called an **affine subspace** or a **linear variety**. If $A \in \mathbb{R}^{1 \times n}$ and is not equal to the zero vector, then $\{x \in \mathbb{R} | Ax = b\}$ is called a **hyperplane**. □

The notion of a hyperplane generalizes the notion of a plane in three dimensions: a hyperplane has exactly one less "dimension" than the space $\mathbb{R}^n$ in which it is embedded. A hyperplane in $\mathbb{R}^n$ divides $\mathbb{R}^n$ into two half-spaces. The boundary of each half-space is the hyperplane.

Definition A.53 In describing matrices, we will mention the:

- **upper triangle**,
- **diagonal**, and

- **lower triangle**.

Some authors use “upper triangle” to refer to both the entries above the diagonal as well as on the diagonal. In this book we will use upper triangle to refer to only the entries above the diagonal. Similarly, we will use “lower triangle” to refer to only the entries below the diagonal. By an **upper triangular matrix**, we will mean a matrix that has zeros in its lower triangle. Similarly, a **lower triangular matrix** has zeros in its upper triangle. □

A.6.1.2 Properties

Theorem A.5 *A square matrix $A \in \mathbb{R}^{n \times n}$ that is singular has the property that there exists a non-zero value of x such that $Ax = \mathbf{0}$. That is, the null space of a singular matrix contains elements besides the zero vector.*

Proof See [55, appendix]. □

Theorem A.6 *Let $A \in \mathbb{R}^{n \times n}$. Suppose that $B \in \mathbb{R}^{n \times n}$ satisfies $AB = \mathbf{I}$. Then $B = A^{-1}$ and $BA = \mathbf{I}$. Similarly, if $B \in \mathbb{R}^{n \times n}$ satisfies $BA = \mathbf{I}$, then $B = A^{-1}$ and $AB = \mathbf{I}$.*

Proof See [55, appendix]. □

In general, for two arbitrary matrices A and B, it is *not* usually the case that $AB = BA$. In the special case that $B = A^{-1}$, this relationship, called **commutativity**, does hold.

A.6.2 Linearly independent columns and rows

Definition A.54 Let $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, $g : \mathbb{R}^n \to \mathbb{R}^m$. Then:

- the column vector Ax is called a **linear combination** of the columns of A,
- the row vector $y^\dagger A$ is called a **linear combination** of the rows of A,
- the function $y^\dagger g : \mathbb{R}^n \to \mathbb{R}$ is called a **linear combination** of the entries of g,
- the equation $y^\dagger g(x) = 0$ is called a **linear combination** of the equations $g(x) = \mathbf{0}$.

□

Definition A.55 A matrix $A \in \mathbb{R}^{m \times n}$ has **linearly independent columns** if:

$$\forall x \in \mathbb{R}^n, (Ax = \mathbf{0}) \Rightarrow (x = \mathbf{0}).$$

It has **linearly independent rows** if:

$$\forall y \in \mathbb{R}^m, (y^\dagger A = \mathbf{0}) \Rightarrow (y = \mathbf{0}).$$

If the matrix does not have linearly independent rows then we can write one of the rows as a linear combination of the others and we say that the rows are **linearly dependent**. If the matrix does not have linearly independent columns then we can write one of the columns as a linear combination of the others and we say that the columns are **linearly dependent**.
□

If A has linearly independent columns then $A^\dagger$ has linearly independent rows. If A has linearly independent rows then $A^\dagger$ has linearly independent columns. In the case that $m = n$ then having linearly independent columns is equivalent to the matrix being non-singular and is equivalent to the matrix having linearly independent rows. If the null space of A has elements besides $\mathbf{0}$ then the columns of A are not linearly independent and vice versa as Exercise A.16 shows.

Definition A.56 A **basis** for a vector subspace is a linearly independent set of vectors such that all the elements of the vector subspace can be expressed as a linear combination of the vectors in the basis. □

If the columns of a matrix A are linearly independent, then the columns form a basis for the range space of A. For example, $\mathbf{I} \in \mathbb{R}^{n \times n}$ has linearly independent columns and the vectors $\{\mathbf{I}_1, \dots, \mathbf{I}_n\}$ are a basis for $\mathbb{R}^n$, which is the range space of $\mathbf{I}$.

Definition A.57 Consider a matrix $A \in \mathbb{R}^{m \times n}$. We define:

- a **row sub-matrix** to be a matrix obtained from A by deleting some of its rows, and
- a **column sub-matrix** to be a matrix obtained from A by deleting some of its columns.

The **row rank** of A is the number of rows in the largest row sub-matrix of A that has linearly independent rows. The **column rank** of A is the number of columns in the largest column sub-matrix of A that has linearly independent columns.

A matrix $A \in \mathbb{R}^{m \times n}$ has full row rank if its row rank is equal to m. It has full column rank if its column rank is equal to n. A square matrix $A \in \mathbb{R}^{n \times n}$ is invertible if and only if it has full row rank and if and only if it has full column rank.

□

A.6.3 Positive definite and positive semi-definite matrices

Definition A.58 A matrix $Q \in \mathbb{R}^{n \times n}$ is **positive definite** if:

$$\forall x \in \mathbb{R}^n, (x \neq \mathbf{0}) \Rightarrow (x^\dagger Q x > 0).$$

A matrix $Q \in \mathbb{R}^{n \times n}$ is **negative definite** if $(-Q)$ is positive definite. □

Definition A.59 A matrix $Q \in \mathbb{R}^{n \times n}$ is **positive semi-definite** if:

$$\forall x \in \mathbb{R}^n, x^\dagger Q x \geq 0.$$

A matrix $Q \in \mathbb{R}^{n \times n}$ is **negative semi-definite** if $(-Q)$ is positive semi-definite. □

A.6.4 Positive definiteness on a subspace

Definition A.60 A matrix $Q \in \mathbb{R}^{n\times n}$ is **positive definite on the null space** $\{x \in \mathbb{R}^n | Ax = \mathbf{0}\}$ if:

$$\forall x \in \mathbb{R}^n, (Ax = \mathbf{0} \text{ and } x \neq \mathbf{0}) \Rightarrow (x^\dagger Qx > 0).$$

□

Definition A.61 A matrix $Q \in \mathbb{R}^{n\times n}$ is **positive semi-definite on the null space** $\{x \in \mathbb{R}^n | Ax = \mathbf{0}\}$ if:

$$\forall x \in \mathbb{R}^n, (Ax = \mathbf{0}) \Rightarrow (x^\dagger Qx \geq 0).$$

□

A.7 Special results

In this section we present some special results.

A.7.1 Weierstrass accumulation principle

Although, in general, sequences may or may not converge, we have the following.

Theorem A.7 *Suppose that the sequence $\{x^{(\nu)}\}_{\nu=0}^{\infty}$ is bounded. (See Definition A.46.) Then it has a convergent sub-sequence. (See Definitions A.18 and A.34.)*

Proof See [111, corollary of theorem 2 of chapter 21]. □

A.7.2 l'Hôpital's rule

In some cases, limits involving ratios can be calculated using **l'Hôpital's rule**.

Theorem A.8 *Let $f, g : \mathbb{R} \to \mathbb{R}$ be differentiable and suppose that $\lim_{x\to 0} f(x) = \lim_{x\to 0} g(x) = 0$. Then:*

$$\lim_{x\to 0} \frac{f(x)}{g(x)} = \lim_{x\to 0} \frac{\frac{df}{dx}(x)}{\frac{dg}{dx}(x)},$$

assuming that the limit on the right-hand side exists.

Proof See [111, theorem 9 of chapter 11]. □

A.7.3 Implicit function theorem

In discussing sensitivity analysis, we are interested in how an optimal solution varies with the values of a parameter.

Theorem A.9 *Let $g : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}^n$ be partially differentiable with continuous partial derivatives. Consider solutions of the equations $g(x; \chi) = \mathbf{0}$, where $\chi \in \mathbb{R}^s$ is a parameter. Suppose that $x^{\star\star} \in \mathbb{R}^n$ is a solution, satisfying:*

$$g(x^{\star\star}; \mathbf{0}) = \mathbf{0}.$$

We call $x = x^{\star\star}$ the base-case solution and $\chi = \mathbf{0}$ the base-case parameters. Define the (parameterized) Jacobian $J : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}^{n \times n}$ by:

$$\forall x \in \mathbb{R}^n, \forall \chi \in \mathbb{R}^s, J(x; \chi) = \frac{\partial g}{\partial x}(x; \chi).$$

Suppose that $J(x^{\star\star}; \mathbf{0})$ is non-singular. Then, there exists a neighborhood $\mathbb{P}$ of $\chi = \mathbf{0}$ and a partially differentiable function $x^\star : \mathbb{R}^s \to \mathbb{R}^n$ with continuous partial derivatives such that:

- *$x^\star(\mathbf{0}) = x^{\star\star}$ is equal to the base-case solution,*
- *$x^\star$ satisfies:*

$$\forall \chi \in \mathbb{P}, g(x^\star(\chi); \chi) = \mathbf{0},$$

  and
- *the sensitivity of $x^\star$ to variation of the parameters satisfies:*

$$\frac{\partial x^\star}{\partial \chi}(\mathbf{0}) = -[J(x^\star(\mathbf{0}); \mathbf{0})]^{-1} K(x^\star(\mathbf{0}); \mathbf{0}),$$

  where $K : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}^{n \times s}$ is defined by:

$$\forall x \in \mathbb{R}^n, \forall \chi \in \mathbb{R}^s, K(x; \chi) = \frac{\partial g}{\partial \chi}(x; \chi).$$

Proof See [70, section A.6][72, section 4.4]. □

The most straightforward application of the implicit function theorem is in calculating the sensitivity to χ of the solution of simultaneous equations evaluated at the base-case. This is considered in Section 7.5.

Since the base-case solution $x^{\star\star}$ in Theorem A.9 is equal to $x^\star(\mathbf{0})$, we will usually abuse notation somewhat and write $x^\star$ for both the base-case solution and *also* for the function that represents the dependence of the solution on χ. That is, whether the symbol $x^\star$ stands for a particular vector value or for a vector function will depend on context. Since we are usually only interested in the base-case solution $x^\star$ and its sensitivity at the base-case, $\frac{\partial x^\star}{\partial \chi}(\mathbf{0})$, this will not be ambiguous.

A.7.4 Inverse function theorem

A related result is called the inverse function theorem. It allows us to “invert” a function. See Exercise A.19 and [72, section 4.4] for details.

Exercises

Types of functions

A.1 In the following, $Q \in \mathbb{R}^{n \times n}$ is not necessarily symmetric. Define $Q^\flat = \frac{1}{2}(Q + Q^\dagger)$.

(i) Show that $Q^\flat$ is symmetric.
(ii) Show that $\forall x$, $\frac{1}{2}x^\dagger Q x = \frac{1}{2}x^\dagger Q^\flat x$.
(iii) Show that (Q is positive semi-definite) $\Leftrightarrow$ ($Q^\flat$ is positive semi-definite).
(iv) Show that (Q is positive definite) $\Leftrightarrow$ ($Q^\flat$ is positive definite).

Norms

A.2 In this exercise we consider several norms.

(i) Prove that the L_1 norm $\|\bullet\|_1$ satisfies the definition of a norm.
(ii) Prove that infinity norm $\|\bullet\|_\infty$ satisfies the definition of a norm.
(iii) Show that on $\mathbb{R}^1$ that $\|\bullet\|_1 = \|\bullet\|_2 = \|\bullet\|_\infty$.
(iv) Show that on $\mathbb{R}^n$ each of these three norms is bounded above and below by some constant multiple of the others. Calculate all six constants relating the norms. (Each constant depends on n.)

A.3 Use the triangle inequality (and any other properties of norms that you might need) to prove that for any norm:

$$\|x + y\| \geq \|x\| - \|y\| .$$

A.4 Show that the induced matrix norm in Definition A.30 satisfies Definition A.29 of a matrix norm.

A.5 Consider the matrix $A \in \mathbb{R}^{2\times 2}$ and the L_2 norm $\|\bullet\|_2$. Calculate the value of the induced matrix norm $\|A\|$ for:

(i) $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$,
(ii) $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$,
(iii) $A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$.

Limits

A.6 Do the following sequences $\{x^{(\nu)}\}_{\nu=0}^{\infty}$ have any accumulation points? For each accumulation point, specify a convergent sub-sequence and its accumulation point.

(i) $\forall \nu \in \mathbb{Z}_+, x^{(\nu)} = \begin{cases} (\nu)^2, & \text{if } \nu \text{ is odd,} \\ 1/(\nu+1), & \text{if } \nu \text{ is even.} \end{cases}$

(ii) $\forall \nu \in \mathbb{Z}_+, x^{(\nu)} = \begin{cases} 1, & \text{if } \nu \text{ is divisible by 4,} \\ 1/\nu, & \text{if } \nu \text{ has remainder 1 after division by 4,} \\ (\nu)^2, & \text{if } \nu \text{ has remainder 2 after division by 4,} \\ -1/\nu, & \text{if } \nu \text{ has remainder 3 after division by 4,} \end{cases}$

(iii) $\forall \nu \in \mathbb{Z}_+, x^{(\nu)} = \begin{cases} \begin{bmatrix} (\nu)^2 \\ 1/\nu \end{bmatrix}, & \text{if } \nu \text{ is odd,} \\ \begin{bmatrix} 1/(\nu+1) \\ (\nu+1)^2 \end{bmatrix}, & \text{if } \nu \text{ is even.} \end{cases}$

A.7 Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by:

$$\forall x \in \mathbb{R},\ f(x) = \begin{cases} 1, & \text{if } x = 0, \\ x, & \text{if } x \neq 0. \end{cases}$$

Show that f is not continuous at $x = 0$. Use the $\|\bullet\|_1$ norm.

A.8 Show that a norm on $\mathbb{R}^n$ is a continuous function. (Hint: Notice that $\|\bullet\| : \mathbb{R}^n \to \mathbb{R}$, so you must define norms on $\mathbb{R}^n$ and on $\mathbb{R}$. Which norms should they be to make your work easy?)

A.9 ([72, example 2 of appendix A].) Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by:

$$\forall x \in \mathbb{R}^2,\ f(x) = \begin{cases} \frac{x_1 x_2}{\sqrt{(x_1)^2+(x_2)^2}}, & \text{if } x \neq \mathbf{0}, \\ 0, & \text{if } x = \mathbf{0}. \end{cases}$$

(i) Sketch the function.
(ii) Show that f is partially differentiable at each $x \in \mathbb{R}^2$.
(iii) Show that the partial derivatives are not continuous at $x^\star = \mathbf{0}$.
(iv) Let $\Delta x = \mathbf{1} \in \mathbb{R}^2$ and define the function $\phi : \mathbb{R} \to \mathbb{R}$ by:

$$\forall t \in \mathbb{R}, \phi(t) = f(t\Delta x).$$

Is ϕ continuous at $t = 0$?
(v) For ϕ defined in the previous part, is ϕ differentiable at $t = 0$?

A.10 In this exercise we consider quadratic functions.

(i) Let $Q \in \mathbb{R}^{n\times n}$ be symmetric. Show that the Hessian of f defined in (A.1) is given by Q.
(ii) Suppose that $Q \in \mathbb{R}^{n\times n}$ is not symmetric. What is the Hessian of f defined in (A.1)?

A.11 Let $f : \mathbb{R} \to \mathbb{R}_+$ be continuous and let $a, b \in \mathbb{R}$.

(i) Prove that:

$$\int_{t=a}^{b} f(t)\,dt \geq 0.$$

(ii) Now suppose that $f(t)$ is strictly positive for $a < t < b$. Prove that:

$$\int_{t=a}^{b} f(t)\,dt > 0.$$

Sets

A.12 In this exercise we consider open and closed balls.

(i) Prove that a closed ball is a closed set. Make sure that your proof applies the definitions carefully.
(ii) Prove that an open ball is not a closed set.
(iii) What points would have to be added to the open ball to make it closed? Specify the smallest set of added points that would make the open ball into a closed set.
(iv) Prove that an open ball is an open set.

A.13 Show that the intersection of two closed sets is closed.

A.14 In this exercise we consider sets defined in terms of functions.

(i) Let $g : \mathbb{R}^n \to \mathbb{R}^m$ be continuous. Show that $\mathbb{S} = \{x \in \mathbb{R}^n | g(x) = \mathbf{0}\}$ is closed.
(ii) Let $h : \mathbb{R}^n \to \mathbb{R}^r$ be continuous. Show that $\mathbb{S} = \{x \in \mathbb{R}^n | h(x) \leq \mathbf{0}\}$ is closed.
(iii) Let $g : \mathbb{R}^n \to \mathbb{R}^m$ and $h : \mathbb{R}^n \to \mathbb{R}^r$ be continuous. Show that $\mathbb{S} = \{x \in \mathbb{R}^n | g(x) = \mathbf{0}, h(x) \leq \mathbf{0}\}$ is closed.

A.15 Suppose that $h : \mathbb{R}^n \to \mathbb{R}^r$ is continuous and consider the sets $\mathbb{S} = \{\mathbb{R}^n | h(x) \leq \mathbf{0}\}$, $\hat{\mathbb{S}} = \{x \in \mathbb{R}^n | h(x) \leq \mathbf{0}$ and, for at least one ℓ, $h_\ell(x) = 0\}$, and $\underline{\mathbb{S}} = \{x \in \mathbb{R}^n | h(x) < \mathbf{0}\}$.

(i) Suppose that each element of $\mathbb{S}$ is a regular point of the constraints $h(x) \leq \mathbf{0}$. (See Definition 19.1.) Show that the interior of $\mathbb{S}$ is $\underline{\mathbb{S}}$ and that the boundary of $\mathbb{S}$ is $\hat{\mathbb{S}}$.
(ii) Suppose that h is a convex function (see Definition 2.16) and that $\underline{\mathbb{S}} \neq \emptyset$. Show that the interior of $\mathbb{S}$ is $\underline{\mathbb{S}}$ and that the boundary of $\mathbb{S}$ is $\hat{\mathbb{S}}$.
(iii) Show by an example that if h is not continuous then the boundary of $\mathbb{S}$ is not necessarily $\hat{\mathbb{S}}$.
(iv) Show by an example that if h is continuous but some elements of $\mathbb{S}$ are not regular points of the constraints $h(x) \leq \mathbf{0}$ then the interior of $\mathbb{S}$ is not necessarily $\underline{\mathbb{S}}$.
(v) Show by an example that if h is continuous and convex but $\underline{\mathbb{S}} = \emptyset$ then the interior of $\mathbb{S}$ is not necessarily $\underline{\mathbb{S}}$.

Properties of matrices

A.16 Let $A \in \mathbb{R}^{m \times n}$. Show that the columns of A are not linearly independent if and only if the null space of A contains vectors besides $\mathbf{0}$.

A.17 Do the following have linearly independent columns? What is the column rank of each matrix?

(i) $A = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

(ii) $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

(iii) $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$.

(iv) $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.

(v) $A = \begin{bmatrix} -1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$.

(vi) $A = \begin{bmatrix} 1 & 1 & 3 \\ 2 & 2 & 4 \end{bmatrix}$.

(vii) $A = \begin{bmatrix} 1 & 2 & 6 \\ 2 & 3 & 10 \end{bmatrix}$.

(viii) $A = \begin{bmatrix} 0 & 1 \\ 0 & 2 \\ 0 & 3 \end{bmatrix}$.

A.18 Let $Q \in \mathbb{R}^{n \times n}$ and $A \in \mathbb{R}^{m \times n}$.

(i) Suppose that there exists $\Pi \in \mathbb{R}_+$ such that $Q + \Pi A^\dagger A$ is positive definite. Show that Q is positive definite on the null space $\mathcal{N}(A) = \{\Delta x \in \mathbb{R}^n | A\Delta x = \mathbf{0}\}$.

(ii) Suppose that Q is positive definite on the null space $\mathcal{N}(A) = \{\Delta x \in \mathbb{R}^n | A\Delta x = \mathbf{0}\}$. Show that there exists $\Pi \in \mathbb{R}_+$ such that $Q + \Pi A^\dagger A$ is positive definite. (Hint: Prove by contradiction. Suppose that for each $\nu \in \mathbb{Z}_+$ there is $x^{(\nu)}$ such that $\|x^{(\nu)}\| = 1$ and $[x^{(\nu)}]^\dagger (Q + \nu A^\dagger A) x^{(\nu)} \leq 0$. Apply Theorem A.7 to find a convergent sub-sequence of $\{x^{(\nu)}\}_{\nu=1}^{\infty}$.)

Special results

A.19 Let $h : \mathbb{R}^n \to \mathbb{R}^n$ be partially differentiable with continuous partial derivatives and $x^{\star\star} \in \mathbb{R}^n$. Suppose that $h(x^{\star\star}) = \mathbf{0}$ and that $\frac{\partial h}{\partial x}(x^{\star\star})$ is non-singular. Use the implicit function theorem, Theorem A.9, to show that, in a neighborhood of $\chi = \mathbf{0}$, there exists an inverse function $x^\star : \mathbb{R}^n \to \mathbb{R}^n$ to h. In particular, show that there exists a neighborhood $\mathbb{P}$ of $\chi = \mathbf{0}$ and a partially differentiable function $x^\star : \mathbb{R}^n \to \mathbb{R}^n$ with continuous partial derivatives such that:

- $x^\star(\mathbf{0}) = x^{\star\star}$,
- $x^\star$ satisfies:

$$\forall \chi \in \mathbb{P}, h(x^\star(\chi)) = \chi,$$

  and
- the sensitivity of $x^\star$ to variation of χ satisfies:

$$\frac{\partial x^\star}{\partial \chi}(\mathbf{0}) = \left[\frac{\partial h}{\partial x}(x^{\star\star})\right]^{-1}.$$

(Hint: Define $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ by $\forall x \in \mathbb{R}^n, \forall \chi \in \mathbb{R}^n, g(x; \chi) = h(x) - \chi$.)

Appendix B
Proofs of theorems

B.1 Problems, algorithms, and solutions

Theorem 2.6 We follow the proof of [70, proposition 4, section 6.4].

$\Rightarrow$ Suppose that f is convex. Let $x, x' \in \mathbb{S}$ be given. Then, by definition,

$$\forall t \in [0, 1],\ f(x' + t[x - x']) \leq f(x') + t[f(x) - f(x')].$$

Re-arranging and dividing through by t for $0 < t \leq 1$, we obtain:

$$\forall t \in (0, 1],\ \frac{f(x' + t[x - x']) - f(x')}{t} \leq f(x) - f(x'). \tag{B.1}$$

To interpret (B.1), consider a line interpolating f between x' and x as shown in Figure B.1. This line has slope:

$$\frac{f(x) - f(x')}{\|x - x'\|_2},$$

and is illustrated with the dashed line in Figure B.1. Now consider a line interpolating f between x' and $x' + t[x - x']$. This line has slope:

$$\frac{f(x' + t[x - x']) - f(x')}{t\ \|x - x'\|_2},$$

and is illustrated with the dash-dotted line in Figure B.1. Equation (B.1) shows that the slope of the dash-dotted line is no greater than the slope of the dashed line. This is true for each value of t in the range $0 < t \leq 1$. The situation is illustrated in Figure B.1.

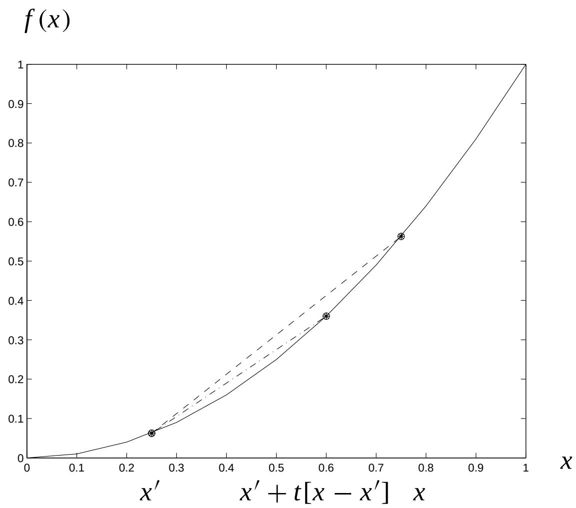

Fig. B.1. Graphical illustration of inequality (B.1) in Theorem 2.6. The line interpolating f between x' and x is shown dashed, while the line interpolating f between x' and $x'+t[x-x']$ is shown dash-dotted.

Moreover, since f is partially differentiable with continuous partial derivatives,

$$\begin{aligned}
&\nabla f(x')^{\dagger}(x - x') \\
&= \lim_{t\to 0} \frac{f(x' + t[x - x']) - f(x')}{t}, \\
&\quad \text{by definition of the partial derivative (see Definition A.36),} \\
&\quad \text{and of the directional derivative (see Definition A.37),} \\
&\leq \lim_{t\to 0}[f(x) - f(x')], \text{ by (B.1), on replacing } (f(x' + t[x - x']) - f(x'))/t \\
&\quad \text{with the value } f(x) - f(x'), \text{ which is always greater,} \\
&= f(x) - f(x').
\end{aligned}$$

The result is true for arbitrary $x, x' \in \mathbb{S}$ so that (2.31) holds.

$\Leftarrow$ Conversely, suppose that (2.31) holds. Let $x, x'' \in \mathbb{S}$ and $0 \leq t \leq 1$ be arbitrary. To prove that f is convex, we must show that $f(x + t[x'' - x]) \leq f(x) + t[f(x'') - f(x)]$. Let $x' = x + t[x'' - x]$. Then, equivalently, we must prove that $f(x') \leq f(x) + t[f(x'') - f(x)]$. Now notice that:

$$f(x) + t[f(x'') - f(x)] = [1 - t]f(x) + tf(x''),$$

so that equivalently we must show that:

$$f(x') \leq [1 - t]f(x) + tf(x''). \tag{B.2}$$

By (2.31), since $x, x' \in \mathbb{S}$,

$$f(x) \geq f(x') + \nabla f(x')^{\dagger}(x - x'). \tag{B.3}$$

But by (2.31) applied to x'' and x', (that is, replacing x by x'' in (2.31) and observing that $x'', x' \in \mathbb{S}$),

$$f(x'') \geq f(x') + \nabla f(x')^{\dagger}(x'' - x'). \tag{B.4}$$

Now multiply (B.3) by $[1 - t]$ and multiply (B.4) by t and add the results together to obtain:

$$\begin{aligned}
&[1-t]f(x) + tf(x'') \\
&\quad \geq \quad [1-t]f(x') + [1-t]\nabla f(x')^{\dagger}(x - x') + tf(x') + t\nabla f(x')^{\dagger}(x'' - x'), \\
&\quad = \quad f(x') + \nabla f(x')^{\dagger}[(1-t)(x - x') + t(x'' - x')], \\
&\quad = \quad f(x') + \nabla f(x')^{\dagger}[x + t(x'' - x) - x'], \\
&\quad \qquad \text{on collecting and re-arranging the terms in the square brackets,} \\
&\quad = \quad f(x') + \nabla f(x')^{\dagger}[\mathbf{0}], \text{ by definition of } x', \\
&\quad = \quad f(x'),
\end{aligned}$$

which is (B.2). □

Theorem 2.7 By Theorem 2.6, we must show that (2.31) holds. Let $x, x' \in \mathbb{S}$. For $0 \leq t \leq 1$ we have that $(x' + t[x - x']) \in \mathbb{S}$ since $\mathbb{S}$ is convex. Define $\phi : [0, 1] \to \mathbb{R}$ by $\forall t \in [0, 1], \phi(t) = f(x' + t[x - x'])$. Notice that:

$$\begin{aligned}
\phi(0) &= f(x'), & \text{(B.5)} \\
\phi(1) &= f(x). & \text{(B.6)}
\end{aligned}$$

Taking derivatives:

$$\begin{aligned}
\frac{d\phi}{dt}(t) &= \nabla f(x' + t[x - x'])^{\dagger}(x - x'), \text{ by the chain rule [72, section 2.4]}, \\
\frac{d\phi}{dt}(0) &= \nabla f(x')^{\dagger}(x - x'), \text{ evaluating the previous expression at } t = 0, \quad \text{(B.7)} \\
\frac{d^2\phi}{dt^2}(t) &= (x - x')^{\dagger}\nabla^2 f(x' + t[x - x'])(x - x'), \\
&\geq 0, \text{ for } 0 \leq t \leq 1 \text{ since } \nabla^2 f(x' + t[x - x']) \text{ is positive semi-definite.} \\
&& \text{(B.8)}
\end{aligned}$$

By (B.5), (B.6), and (B.7), the condition (2.31) is equivalent to $\phi(1) \geq \phi(0) + \frac{d\phi}{dt}(0)$. We have:

$$\begin{aligned}
\phi(1) &= \phi(0) + \int_{t=0}^{1} \frac{d\phi}{dt}(t)\,dt, \\
&\quad \text{by the fundamental theorem of integral calculus applied to } \phi, \\
&\qquad \text{(see Theorem A.2 in Section A.4.4.1 of Appendix A)}, \\
&= \phi(0) + \int_{t=0}^{1} \left[\frac{d\phi}{dt}(0) + \int_{t'=0}^{t} \frac{d^2\phi}{dt^2}(t')\,dt'\right] dt, \\
&\quad \text{by the fundamental theorem of integral calculus applied to } \frac{d\phi}{dt}, \\
&\qquad \text{(see Theorem A.2 in Section A.4.4.1 of Appendix A)}, \\
&= \phi(0) + \frac{d\phi}{dt}(0) + \int_{t=0}^{1} \int_{t'=0}^{t} \frac{d^2\phi}{dt^2}(t')\,dt'\,dt, \qquad \text{(B.9)} \\
&\quad \text{evaluating the integral of the first term in the integrand}, \\
&\geq \phi(0) + \frac{d\phi}{dt}(0), \\
&\quad \text{since the integrand is non-negative everywhere by (B.8)}, \\
&\qquad \text{(see Theorem A.3 in Section A.4.4.2 of Appendix A)}.
\end{aligned}$$

This is the result we were trying to prove. A similar analysis applies if $\nabla^2 f$ is positive definite, where we note that continuity and positive definiteness of the Hessian implies that the integrand $\frac{d^2\phi}{dt^2}(t')$ in (B.9) is continuous and strictly positive everywhere. □

B.2 Algorithms for linear simultaneous equations

Lemma 5.1 First notice that the symmetry of A is preserved when we re-order the rows and columns using diagonal pivoting. Therefore, we can assume that A has its rows and columns ordered so that A_{11} is the first pivot. By definition,

$$\forall \ell = 2, \ldots, n, L_{\ell 1} = A_{\ell 1}/A_{11}.$$

Now consider any entry $A_{\ell k}^{(2)}$ with $\ell, k \geq 2$. We have:

$$\begin{aligned}
A_{\ell k}^{(2)} &= A_{\ell k} - L_{\ell 1} A_{1k}, \text{ by definition of } A^{(2)}, \\
&= A_{\ell k} - A_{\ell 1} A_{1k} / A_{11}, \text{ by definition of } L_{\ell 1}. \\
\text{Also, } A_{k\ell}^{(2)} &= A_{k\ell} - L_{k1} A_{1\ell}, \text{ by definition of } A^{(2)}, \\
&= A_{k\ell} - A_{k1} A_{1\ell} / A_{11}, \text{ by definition of } L_{k1}, \\
&= A_{\ell k} - A_{\ell 1} A_{1k} / A_{11}, \text{ by symmetry of } A, \\
&= A_{\ell k}^{(2)}.
\end{aligned}$$

□

Lemma 5.2 Again, we can assume that $A_{jj}^{(j)}$ was used as the pivot. Then,

$$\forall \ell > j, L_{\ell j} = A_{\ell j}^{(j)} / A_{jj}^{(j)},$$

by definition. Consider any entry $A_{\ell k}^{(j+1)}$ with $\ell, k \geq j + 1$. We have

$$\begin{aligned}
A_{\ell k}^{(j+1)} &= A_{\ell k}^{(j)} - L_{\ell j} A_{jk}^{(j)}, \text{ by definition of } A^{(j+1)}, \\
&= A_{\ell k}^{(j)} - A_{\ell j}^{(j)} A_{jk}^{(j)} / A_{jj}^{(j)}, \text{ by definition of } L_{\ell j}. \\
\text{Also, } A_{k\ell}^{(j+1)} &= A_{k\ell}^{(j)} - L_{kj} A_{j\ell}^{(j)}, \text{ by definition of } A^{(j+1)}, \\
&= A_{k\ell}^{(j)} - A_{kj}^{(j)} A_{j\ell}^{(j)} / A_{jj}^{(j)}, \text{ by definition of } L_{kj}, \\
&= A_{\ell k}^{(j)} - A_{\ell j}^{(j)} A_{jk}^{(j)} / A_{jj}^{(j)}, \text{ by symmetry of } A^{(j)}, \\
&= A_{\ell k}^{(j+1)}.
\end{aligned}$$

□

Theorem 5.5 We divide the proof into three parts.

A **is invertible** Suppose that $A \in \mathbb{R}^{n \times n}$ is singular. Then, by Theorem A.5 in Section A.6.1.2 of Appendix A, there exists $x \neq \mathbf{0}$ such that $Ax = \mathbf{0}$. But then $x^\dagger A x = x^\dagger \mathbf{0} = 0$ and so A is not positive definite. This is a contradiction and so A is, in fact, non-singular. (Positive definiteness is a "stronger" condition than being invertible.)

A **is factorizable as** $LDL^\dagger$ We now claim that we can use the standard pivot $A_{\ell\ell}^{(\ell)}$ at each stage of the factorization algorithm to factorize symmetric positive definite A into LU. For suppose not. That is, suppose that the factorization failed at, say, stage ℓ. By this we mean that factorization using the standard pivot was successful for stages $1, \dots, (\ell - 1)$, but we found that at stage ℓ, $A_{\ell\ell}^{(\ell)} = 0$. (If we find a zero pivot at the first stage, then $\ell = 1$ and we define $A^{(1)} = A$. In this particular case, $A^{(\ell)}_{\ell\ell} = A^{(1)}_{11} = A_{11} = 0$.)

Let L' be the product of the inverses of the matrices $M^{(1)}, \dots, M^{(\ell-1)}$ defined in the factorization algorithm in Section 5.3.2. (If $\ell = 1$, then define $L' = \mathbf{I}$.) Notice that L' is lower triangular with ones on its diagonal. We have $A = L'A^{(\ell)}$.

Consider the top left-hand $\ell \times \ell$ submatrices of A, L', and $A^{(\ell)}$ and write $\hat{A}$, $\hat{L}'$, and $\hat{A}^{(\ell)}$, respectively, for these three $\ell \times \ell$ submatrices. By construction, the matrix $\hat{L}'$ is lower triangular, while $\hat{A}^{(\ell)}$ is upper triangular. Let us use the symbol $\bullet$ to stand for blocks of a matrix that have unknown and possibly non-zero entries. Then we can write:

$$\begin{aligned}
A &= \begin{bmatrix} \hat{A} & \bullet \\ \bullet & \bullet \end{bmatrix}, \text{ by definition of } \hat{A}, \\
&= L'A^{(\ell)}, \text{ by construction}, \\
&= \begin{bmatrix} \hat{L}' & \mathbf{0} \\ \bullet & \bullet \end{bmatrix} \begin{bmatrix} \hat{A}^{(\ell)} & \bullet \\ \bullet & \bullet \end{bmatrix}, \\
&\quad \text{by definition of } L', \hat{L}', \text{ and } \hat{A}^{(\ell)} \text{ and since } L' \text{ is lower triangular}, \\
&= \begin{bmatrix} \hat{L}'\hat{A}^{(\ell)} & \bullet \\ \bullet & \bullet \end{bmatrix}, \text{ on multiplying.}
\end{aligned}$$

Therefore, $\hat{A} = \hat{L}'\hat{A}^{(\ell)}$. For example, if we encounter a zero pivot at the first stage, then $\hat{A} = [A_{11}] = [0]$, $\hat{L}' = [1]$, $\hat{A}^{(1)} = [0]$, and $[0] = [1][0]$. Since $\hat{A}^{(\ell)}$ is upper triangular, if we let $\hat{U}' = \hat{A}^{(\ell)}$ then the LU factorization of $\hat{A}$ is given by $\hat{A} = \hat{L}'\hat{U}'$.

Furthermore, $\hat{A}$ is symmetric and has been factorized into $\hat{L}'\hat{U}'$ using diagonal pivots. Recall that from Corollary 5.3 and the discussion in Section 5.4.4 that if we define $\hat{D}$ to be diagonal with diagonal entries equal to the diagonal of $\hat{U}' = \hat{A}^{(\ell)}$, then we can factor $\hat{A}$ as $\hat{L}'\hat{D}[\hat{L}']^{\dagger}$, where $\hat{L}'$ is lower triangular with ones on the diagonal and $\hat{D}$ is diagonal. Since $A^{(\ell)}_{\ell\ell} = 0$, we have that $D_{\ell\ell} = 0$. By Lemma 5.4 applied to $\hat{A}$, $\hat{A}$ is not positive definite, since the entry $D_{\ell\ell}$ is not positive.

In summary, if the factorization fails at stage ℓ then $\hat{A}$, the top left-hand $\ell \times \ell$ sub-matrix of A, is not positive definite.

But let $\hat{x} \in \mathbb{R}^{\ell}$ be given and define $x = \begin{bmatrix} \hat{x} \\ \mathbf{0} \end{bmatrix} \in \mathbb{R}^n$. We have:

$$\begin{aligned}
\hat{x} \neq \mathbf{0} \quad &\Rightarrow \quad x = \begin{bmatrix} \hat{x} \\ \mathbf{0} \end{bmatrix} \neq \mathbf{0}, \\
&\Rightarrow \quad \hat{x}^{\dagger}\hat{A}\hat{x} = \begin{bmatrix} \hat{x} \\ \mathbf{0} \end{bmatrix}^{\dagger} A \begin{bmatrix} \hat{x} \\ \mathbf{0} \end{bmatrix} = x^{\dagger}Ax > 0, \\
&\qquad \text{since } A \text{ is positive definite by hypothesis.}
\end{aligned}$$

That is, $\hat{x} \neq \mathbf{0} \Rightarrow \hat{x}^\dagger \hat{A} \hat{x} > 0$. But this is true for any $\hat{x} \neq \mathbf{0}$ and so $\hat{A}$ is positive definite. This contradicts the earlier result, so that in fact we could not have encountered a zero pivot at stage ℓ. Therefore, we can successfully use the algorithm to factorize A as LU. But since A is symmetric, by defining D to be diagonal with diagonal entries equal to the diagonal of U, we can factorize A as $A = LDL^\dagger$.

In conclusion, A can be factorized into $LDL^\dagger$ with L lower triangular having ones on its diagonal and D diagonal. By Lemma 5.4, since A is positive definite, D has strictly positive diagonal entries.

A^{-1} **is positive definite** As we did in the proof of Lemma 5.4, let $D^{\frac{1}{2}}$ be diagonal with diagonal entries equal to the square roots of the corresponding diagonal entries of D. Then:

$$\begin{aligned}
A^{-1} &= [LDL^\dagger]^{-1}, \text{ by assumption on } A, \\
&= [L^\dagger]^{-1} D^{-1} L^{-1}, \text{ recalling from Section 5.3.2 that } L \text{ is invertible}, \\
&= [L^{-1}]^\dagger D^{-1} L^{-1}, \text{ since } [L^{-1}]^\dagger = [L^\dagger]^{-1}, \text{(see Exercise 5.18)}, \\
&= [L^{-1}]^\dagger \left[D^{\frac{1}{2}}\right]^{-1} \left[D^{\frac{1}{2}}\right]^{-1} L^{-1}, \text{ by definition of } D^{\frac{1}{2}}.
\end{aligned}$$

Let $x \neq \mathbf{0}$ be given. Note that $\left[D^{\frac{1}{2}}\right]^{-1} L^{-1} x \neq \mathbf{0}$ (for else $x = LD^{\frac{1}{2}}\mathbf{0} = \mathbf{0}$.) But this means that $x^\dagger A^{-1} x = \left\| \left[D^{\frac{1}{2}}\right]^{-1} L^{-1} x \right\|_2^2 > 0$, by Property (ii) of norms, so that A^{-1} is positive definite. □

Lemma 5.6 To calculate $A^{(j+1)}$, using $A_{jj}^{(j)}$ as pivot, we apply (5.11) to calculate:

$$A_{\ell k}^{(j+1)} = A_{\ell k}^{(j)} - L_{\ell j} A_{jk}^{(j)}, \; j < \ell \leq n, \; j < k \leq n.$$

The number of fill-ins is equal to the number of times that $A_{\ell k}^{(j)} = 0$, yet $L_{\ell j} A_{jk}^{(j)} \neq 0$, so that $A_{\ell k}^{(j+1)} \neq 0$. Define $\mathbf{I}^{(j)} \in \mathbb{R}^{n \times n}$ by:

$$\forall \ell, k, \mathbf{I}_{\ell k}^{(j)} = \begin{cases} 0, & \text{if } A_{\ell k}^{(j)} = 0, \\ 1, & \text{if } A_{\ell k}^{(j)} \neq 0. \end{cases}$$

Then a fill-in is created at the ℓk-th entry if:

(i) $A_{\ell k}^{(j)} = 0$; that is, $\mathbf{I}_{\ell k}^{(j)} = 0$,
(ii) $L_{\ell j} = A_{\ell j}^{(j)} / A_{jj}^{(j)} \neq 0$; that is, $A_{\ell j}^{(j)} \neq 0$ and $\mathbf{I}_{\ell j}^{(j)} = 1$, and
(iii) $A_{jk}^{(j)} \neq 0$; that is, $\mathbf{I}_{jk}^{(j)} = 1$.

Therefore, a fill-in occurs at the ℓk-th entry if and only if $(1 - \mathbf{I}_{\ell k}^{(j)}) \mathbf{I}_{\ell j}^{(j)} \mathbf{I}_{jk}^{(j)} = 1$. If a fill-in does not occur, then $(1 - \mathbf{I}_{\ell k}^{(j)}) \mathbf{I}_{\ell j}^{(j)} \mathbf{I}_{jk}^{(j)} = 0$.

We defined $N(j)$ to be the number of fill-ins created at stage j due to pivoting on $A_{jj}^{(j)}$ at stage j. We can calculate $N(j)$ by summing $(1-\mathbf{I}_{\ell k}^{(j)})\mathbf{I}_{\ell j}^{(j)}\mathbf{I}_{jk}^{(j)}$ over all the ℓk-th entries that are in rows $j+1$ to n and columns $j+1$ to n. That is:

$$N(j) = \sum_{\substack{j<\ell\leq n\\ j<k\leq n}} (1-\mathbf{I}_{\ell k}^{(j)})\mathbf{I}_{\ell j}^{(j)}\mathbf{I}_{jk}^{(j)}. \tag{B.10}$$

Since A is sparse and we are trying to minimize fill-ins, it is reasonable to assume that $A^{(j)}$ is also sparse. That is, it is rare for $A_{\ell k}^{(j)}$ to be non-zero and we can approximate the sum in (B.10) by neglecting the term $(1-\mathbf{I}_{\ell k}^{(j)})$, since it is usually equal to one. We calculate an upper bound, $\overline{N}(j)$, on the number of fill-ins $N(j)$ by neglecting the factor $(1-\mathbf{I}_{\ell k}^{(j)})$. That is:

$$\begin{aligned}
N(j) &\leq \overline{N}(j), \\
&= \sum_{\substack{j<\ell\leq n\\ j<k\leq n}} \mathbf{I}_{\ell j}^{(j)}\mathbf{I}_{jk}^{(j)}, \\
&= \left(\sum_{j<\ell\leq n} \mathbf{I}_{\ell j}^{(j)}\right)\left(\sum_{j<k\leq n} \mathbf{I}_{jk}^{(j)}\right), \text{ separating out the sums,} \\
&= \left(\sum_{j<k\leq n} \mathbf{I}_{jk}^{(j)}\right)^2, \text{ because } A^{(j)} \text{ is symmetric.}
\end{aligned}$$

The last expression is the square of:

[(the number of non-zero entries in the j-th row of $A^{(j)}$) minus 1].

(Recall that the first $(j-1)$ entries in this row are zero because of earlier stages in the factorization.) □

B.3 Algorithms for non-linear simultaneous equations

Theorem 7.2 We divide the proof into four parts:

(i) proving that $\{x^{(\nu)}\}_{\nu=0}^{\infty}$ is Cauchy and has a limit that is contained in $\mathbb{S}$;
(ii) proving that the limit is a fixed point of Φ;
(iii) proving that the fixed point is unique; and
(iv) proving that the sequence converges to the fixed point according to (7.21).

$\{x^{(\nu)}\}_{\nu=0}^{\infty}$ **is Cauchy and has a limit that is contained in** $\mathbb{S}$ We prove that the sequence of iterates is a Cauchy sequence. By Lemma 7.1, this will establish that the sequence of iterates converges to some point $x^\star$, say. To prove that the sequence is Cauchy requires four main steps, which successively bound the difference between various pairs of iterates.

Step 1: We first bound the norm of the difference between two successive iterates:

$$\begin{aligned}
&\left\| x^{(m+1)} - x^{(m)} \right\| \\
&\quad = \left\| \Phi(x^{(m)}) - \Phi(x^{(m-1)}) \right\|, \\
&\qquad \text{by (7.20), the definitions of } x^{(m+1)} \text{ and } x^{(m)}, \\
&\quad \leq L \left\| x^{(m)} - x^{(m-1)} \right\|, \\
&\qquad \text{since } \Phi \text{ is a contraction mapping with Lipschitz constant } L, \\
&\quad \leq (L)^2 \left\| x^{(m-1)} - x^{(m-2)} \right\|, \text{ repeating the same argument,} \\
&\quad \vdots \quad \vdots \\
&\quad \leq (L)^m \left\| x^{(1)} - x^{(0)} \right\|,
\end{aligned} \tag{B.11}$$

repeating the argument a further $(m-2)$ times.

Step 2: We use (B.11) to bound the norm of the difference between the ν-th and 0-th iterate:

$$\begin{aligned}
&\left\| x^{(\nu)} - x^{(0)} \right\| \\
&\quad = \left\| (x^{(\nu)} - x^{(\nu-1)}) + (x^{(\nu-1)} - x^{(\nu-2)}) + \cdots + (x^{(1)} - x^{(0)}) \right\|, \\
&\qquad \text{adding and subtracting terms,} \\
&\quad \leq \left\| x^{(\nu)} - x^{(\nu-1)} \right\| + \left\| x^{(\nu-1)} - x^{(\nu-2)} \right\| + \cdots + \left\| x^{(1)} - x^{(0)} \right\|, \\
&\qquad \text{by the triangle inequality (Property (iii) in Definition A.28} \\
&\qquad \text{of norms in Section A.3.1 of Appendix A) applied repeatedly,} \\
&\quad \leq \sum_{m=0}^{\nu-1} (L)^m \left\| x^{(1)} - x^{(0)} \right\|, \text{ using (B.11) for } m = 0, \ldots, \nu - 1, \\
&\quad = \frac{1 - (L)^\nu}{1 - L} \left\| x^{(1)} - x^{(0)} \right\|, \\
&\qquad \text{using the formula for the sum of a geometric progression,} \\
&\quad \leq \frac{1}{1 - L} \left\| x^{(1)} - x^{(0)} \right\|,
\end{aligned} \tag{B.12}$$

since $0 \leq L < 1$.

Step 3: We use (B.12) to bound the norm of the difference between an arbitrary pair of iterates $x^{(\nu)}$ and $x^{(\nu')}$. Let us first suppose that $\nu \leq \nu'$. Then:

$$\begin{aligned}
&\left\| x^{(\nu')} - x^{(\nu)} \right\| \\
&= \left\| \Phi(x^{(\nu'-1)}) - \Phi(x^{(\nu-1)}) \right\|, \text{ by (7.20)}, \\
&\leq L \left\| x^{(\nu'-1)} - x^{(\nu-1)} \right\|, \text{ since } \Phi \text{ is a contraction mapping}, \\
&\vdots \quad \vdots \\
&\leq (L)^{\nu} \left\| x^{(\nu'-\nu)} - x^{(0)} \right\|, \\
&\quad \text{applying the same argument a further } (\nu - 1) \text{ times}, \\
&\leq \frac{(L)^{\nu}}{1-L} \left\| x^{(1)} - x^{(0)} \right\|, \text{ by (B.12) for the } (\nu' - \nu)\text{-th iterate.}
\end{aligned}$$

Similarly, if $\nu' \leq \nu$ then:

$$\left\| x^{(\nu')} - x^{(\nu)} \right\| \leq \frac{(L)^{\nu'}}{1-L} \left\| x^{(1)} - x^{(0)} \right\|.$$

Combining the two results, we obtain:

$$\left\| x^{(\nu')} - x^{(\nu)} \right\| \leq \frac{(L)^{\min\{\nu,\nu'\}}}{1-L} \left\| x^{(1)} - x^{(0)} \right\|. \tag{B.13}$$

Step 4: We use (B.13) to prove that $\{x^{(\nu)}\}_{\nu=0}^{\infty}$ is a Cauchy sequence. For, let $\epsilon > 0$ be given. We claim that:

$$N = \frac{\ln\left[\epsilon(1-L)/\left\| x^{(1)} - x^{(0)} \right\|\right]}{\ln(L)}$$

will suffice in the definition of a Cauchy sequence. By definition of N, we have that (on re-arranging and taking the exponential of both sides):

$$\frac{(L)^{N}}{1-L} \left\| x^{(1)} - x^{(0)} \right\| = \epsilon,$$

so that for $\nu, \nu' \geq N$ we have that $(L)^{\min\{\nu,\nu'\}} \left\| x^{(1)} - x^{(0)} \right\| /(1-L) \leq \epsilon$. Therefore, by (B.13):

$$\forall \nu, \nu' \geq N, \left\| x^{(\nu')} - x^{(\nu)} \right\| \leq \epsilon,$$

and the sequence is Cauchy. By Lemma 7.1, $\{x^{(\nu)}\}_{\nu=0}^{\infty}$ has a limit, $x^{\star}$, say. But $\mathbb{S}$ is closed and $x^{(\nu)} \in \mathbb{S}, \forall \nu$, so $x^{\star} \in \mathbb{S}$.

$x^\star$ **is a fixed point of** Φ Notice that:

$$\begin{aligned}\left\|\Phi(x^\star) - x^{(\nu)}\right\| &= \left\|\Phi(x^\star) - \Phi(x^{(\nu-1)})\right\|, \text{ by (7.20)},\\ &\leq L\left\|x^\star - x^{(\nu-1)}\right\|, \text{ since } \Phi \text{ is Lipschitz.} \qquad \text{(B.14)}\end{aligned}$$

Taking limits of the left- and right-hand sides of (B.14), and recalling that the norm is a continuous function, (see Exercise A.8,) we obtain that $\|\Phi(x^\star) - x^\star\| \leq L\,\|x^\star - x^\star\| = 0$, so that, by Property (ii) of norms, $\Phi(x^\star) = x^\star$ and $x^\star$ is a fixed point of Φ.

Uniqueness of fixed point Now suppose there are two fixed points $x^\star \neq x^{\star\star}$ of Φ in $\mathbb{S}$. Then,

$$\begin{aligned}\left\|x^\star - x^{\star\star}\right\| &= \left\|\Phi(x^\star) - \Phi(x^{\star\star})\right\|, \text{ since } x^\star \text{ and } x^{\star\star} \text{ are fixed points of } \Phi,\\ &\leq L\left\|x^\star - x^{\star\star}\right\|, \text{ since } \Phi \text{ is Lipschitz},\\ &< \left\|x^\star - x^{\star\star}\right\|, \text{ since } L < 1 \text{ because } x^\star \neq x^{\star\star} \text{ by supposition.}\end{aligned}$$

But this is a contradiction. So, there is exactly one fixed point, $x^\star$, say.

Rate of convergence Now note that:

$$\begin{aligned}&\left\|x^{(\nu)} - x^\star\right\|\\ &\quad= \left\|\Phi(x^{(\nu-1)}) - x^\star\right\|, \text{by (7.20)},\\ &\quad= \left\|\Phi(x^{(\nu-1)}) - \Phi(x^\star)\right\|, \text{ since } x^\star \text{ is a fixed point of } \Phi,\\ &\quad\leq L\left\|x^{(\nu-1)} - x^\star\right\|, \text{by definition of contraction mapping},\\ &\quad\leq L^2\left\|x^{(\nu-2)} - x^\star\right\|, \text{repeating the same argument in the last three lines},\\ &\quad\vdots\\ &\quad\leq L^\nu\left\|x^{(0)} - x^\star\right\|, \text{repeating the argument a further } (\nu-2) \text{ times.}\end{aligned}$$

So as $\nu \to \infty$, $L^\nu \to 0$, and $x^{(\nu)} \to x^\star$. That is, the iterative method (7.20) converges to the unique fixed point of Φ in $\mathbb{S}$. Furthermore, the error improves by a factor L at each iteration, satisfying the bound (7.21). $\square$

Theorem 7.3 We reproduce the proof from [58, section 5.5] and divide it into four parts:

(i) we first prove that the iterates stay in $\mathbb{S} = \left\{x \in \mathbb{R}^n \,\middle|\, \left\|x - x^{(0)}\right\| \leq \rho_-\right\}$;

(ii) we then go on to prove that the chord method iteration defines a contraction mapping on $\mathbb{S}$;

(iii) we then prove that the sequence of iterates converges to a solution $x^\star \in \mathbb{R}^n$ of (7.1) that satisfies the estimate (7.22); and

(iv) finally, we prove that $x^\star$ is the only solution in the open ball of radius ρ_+ about $x^{(0)}$.

The iterates stay in $\mathbb{S}$ Consider the map:

$$\Phi(x) = x - [J(x^{(0)})]^{-1} g(x),$$

which specifies the chord method (7.8)–(7.9) in the form (7.20). We show that Φ maps $\mathbb{S}$ to itself. This requires us to estimate the value of $[J(x^{(0)})]^{-1} g(x)$ in terms of known quantities. We know properties of $[J(x^{(0)})]^{-1} g(x^{(0)})$ and J, so we will express $[J(x^{(0)})]^{-1} g(x)$ in terms of these. This requires five main steps.

Step 1: First:

$$[J(x^{(0)})]^{-1} g(x) = [J(x^{(0)})]^{-1} g(x^{(0)}) + [J(x^{(0)})]^{-1} \times (g(x) - g(x^{(0)})), \tag{B.15}$$

on adding and subtracting $[J(x^{(0)})]^{-1} g(x^{(0)})$.

Step 2: We evaluate $(g(x) - g(x^{(0)}))$, the factor in the second term on the right-hand side of (B.15). Define $\gamma : [0, 1] \to \mathbb{R}^n$ by:

$$\forall t \in [0, 1], \gamma(t) = g(x^{(0)} + t(x - x^{(0)})).$$

Then $\gamma(0) = g(x^{(0)})$, $\gamma(1) = g(x)$, and, by the chain rule [72, section 2.4]),

$$\frac{d\gamma}{dt}(t) = J(x^{(0)} + t(x - x^{(0)})) \times (x - x^{(0)}).$$

Therefore:

$$\begin{aligned} g(x) - g(x^{(0)}) &= \gamma(1) - \gamma(0), \\ &= \int_{t=0}^{t=1} \frac{d\gamma}{dt}(t)\, dt, \\ &\quad \text{by the fundamental theorem of integral calculus} \\ &\quad \text{(Theorem A.2 in Section A.4.4.1 of Appendix A),} \\ &= \int_{t=0}^{t=1} [J(x^{(0)} + t(x - x^{(0)}))] \times (x - x^{(0)})\, dt. \end{aligned} \tag{B.16}$$

Step 3: Substituting (B.16) into (B.15), we obtain:

$$\begin{aligned}
&[J(x^{(0)})]^{-1} g(x) \\
&= [J(x^{(0)})]^{-1} g(x^{(0)}) \\
&\quad + [J(x^{(0)})]^{-1} \int_{t=0}^{t=1} [J(x^{(0)} + t(x - x^{(0)}))] \times (x - x^{(0)})\, dt, \\
&= [J(x^{(0)})]^{-1} g(x^{(0)}) + (x - x^{(0)}) - [J(x^{(0)})]^{-1} J(x^{(0)}) \times (x - x^{(0)}) \\
&\quad + [J(x^{(0)})]^{-1} \int_{t=0}^{t=1} [J(x^{(0)} + t(x - x^{(0)}))] \times (x - x^{(0)})\, dt, \\
&\text{adding and subtracting } (x - x^{(0)}) = [J(x^{(0)})]^{-1} J(x^{(0)}) \times (x - x^{(0)}), \\
&= [J(x^{(0)})]^{-1} g(x^{(0)}) + (x - x^{(0)}) + [J(x^{(0)})]^{-1} \times \\
&\left[\int_{t=0}^{t=1} [J(x^{(0)} + t(x - x^{(0)}))] \times (x - x^{(0)})\, dt - J(x^{(0)}) \times (x - x^{(0)})\right], \\
&\qquad\qquad \text{on re-arranging,} \\
&= [J(x^{(0)})]^{-1} g(x^{(0)}) + (x - x^{(0)}) \\
&+ [J(x^{(0)})]^{-1} \int_{t=0}^{t=1} \left[J(x^{(0)} + t(x - x^{(0)})) - J(x^{(0)})\right] \times (x - x^{(0)})\, dt,
\end{aligned} \tag{B.17}$$

where the last equality holds since the integral of a constant between 0 and 1 is equal to the constant.

Step 4: We have:

$$\begin{aligned}
&\Phi(x) - x^{(0)} \\
&= x - [J(x^{(0)})]^{-1} g(x) - x^{(0)}, \text{ by definition,} \\
&= -[J(x^{(0)})]^{-1} g(x^{(0)}) \\
&\quad - [J(x^{(0)})]^{-1} \int_{t=0}^{t=1} \left[J(x^{(0)} + t(x - x^{(0)})) - J(x^{(0)})\right] \times (x - x^{(0)})\, dt,
\end{aligned}$$

from (B.17).

Taking norms and using the triangle inequality and Lemma A.1 in Section A.3.2 of Appendix A repeatedly, we obtain:

$$\begin{aligned}
&\left\| \Phi(x) - x^{(0)} \right\| \\
&\le \left\| [J(x^{(0)})]^{-1} g(x^{(0)}) \right\| + \left\| [J(x^{(0)})]^{-1} \right\| \times \\
&\quad \int_{t=0}^{t=1} \left\| J(x^{(0)} + t(x - x^{(0)})) - J(x^{(0)}) \right\| \left\| x - x^{(0)} \right\| \, dt,
\end{aligned}$$

$$\leq \quad b + a \int_{t=0}^{t=1} ct \left\| x - x^{(0)} \right\|^2 \, dt, \text{ for } x \in \mathbb{S},$$

by the definitions of a, b, and c, since:

- J is Lipschitz with constant c on $\mathbb{S}$;
- $x^{(0)} + t(x - x^{(0)}) \in \mathbb{S}$; and
- $(x^{(0)} + t(x - x^{(0)})) - x^{(0)} = t(x - x^{(0)})$,

$$\leq \quad b + ac\rho_-^2/2, \tag{B.18}$$

evaluating the integral and noting that $\left\| x - x^{(0)} \right\| \leq \rho_-$.

Step 5: By definition of ρ_-:

$$\begin{aligned} \rho_-^2 &= \frac{1 - 2\sqrt{1 - 2abc} + (1 - 2abc)}{(ac)^2}, \\ &= \frac{2(1 - \sqrt{1 - 2abc}) - 2abc}{(ac)^2}, \\ &= \frac{2\rho_- - 2b}{ac}, \end{aligned}$$

so $b + ac\rho_-^2/2 = \rho_-$. Therefore, by (B.18), $\left\| \Phi(x) - x^{(0)} \right\| \leq \rho_-$ and so Φ maps $\mathbb{S}$ to itself.

Φ is a contraction mapping We now show that Φ is a contraction mapping. This requires three main steps.

Step 1: First:

$$\begin{aligned} \frac{\partial \Phi}{\partial x}(x) &= \mathbf{I} - [J(x^{(0)})]^{-1} J(x), \text{ by definition of } J, \\ &= [J(x^{(0)})]^{-1} \times (J(x^{(0)}) - J(x)). \end{aligned}$$

Therefore, for $x \in \mathbb{S}$:

$$\begin{aligned} \left\| \frac{\partial \Phi}{\partial x}(x) \right\| &\leq \left\| [J(x^{(0)})]^{-1} \right\| \left\| J(x^{(0)}) - J(x) \right\|, \text{ by Lemma A.1}, \\ &\leq ac \left\| x^{(0)} - x \right\|, \text{ by assumption}, \\ &\leq ac\rho_-, \text{ since } \left\| x - x^{(0)} \right\| \leq \rho_-. \end{aligned} \tag{B.19}$$

Step 2: Let $x', x'' \in \mathbb{S}$ and define $\phi : \mathbb{R} \to \mathbb{R}^n$ by $\phi(t) = \Phi(x'' + t(x' - x''))$. Then by the chain rule [72, section 2.4]:

$$\frac{d\phi}{dt}(t) = \frac{\partial \Phi}{\partial x}(x'' + t(x' - x'')) \times (x' - x''),$$

and so:

$$
\begin{aligned}
&\Phi(x') - \Phi(x'') \\
&\quad = \phi(1) - \phi(0), \\
&\quad = \int_{t=0}^{t=1} \frac{d\phi}{dt}(t)\, dt, \\
&\qquad \text{by the fundamental theorem of integral calculus,} \\
&\qquad \text{(see Theorem A.2 in Section A.4.4.1 of Appendix A),} \\
&\quad = \int_{t=0}^{t=1} \left[\frac{\partial \Phi}{\partial x}(x'' + t(x' - x''))\right] \times (x' - x'')\, dt.
\end{aligned}
$$

Therefore, on taking norms and using Lemma A.1:

$$
\begin{aligned}
\left\|\Phi(x') - \Phi(x'')\right\| &\leq \int_{t=0}^{t=1} \left\|\frac{\partial \Phi}{\partial x}(x'' + t(x' - x''))\right\| \left\|x' - x''\right\| dt, \\
&\leq \int_{t=0}^{t=1} ac\rho_- \left\|x' - x''\right\| dt, \\
&\qquad \text{by (B.19) since } x'' + t(x' - x'') \in \mathbb{S}, \\
&= ac\rho_- \left\|x' - x''\right\|.
\end{aligned}
$$

Step 3: By definition, $ac\rho_- = 1 - \sqrt{1 - 2abc} < 1$, so Φ is a contraction with Lipschitz constant $L = ac\rho_- < 1$. Therefore, by Theorem 7.2, there is a unique fixed point $x^\star$ of Φ in $\mathbb{S}$ and, moreover, the chord iteration converges to $x^\star$.

The fixed point $x^\star$ satisfies (7.1) and (7.22) Notice that:

$$
(\Phi(x^\star) = x^\star) \Rightarrow \left([J(x^{(0)})]^{-1} g(x^\star) = \mathbf{0}\right) \Rightarrow (g(x^\star) = \mathbf{0}),
$$

so that $x^\star$ is a solution of (7.1). Furthermore, since $x^\star \in \mathbb{S}$, we have that:

$$
\left\|x^{(0)} - x^\star\right\| \leq \rho_-.
$$

Substituting this and the Lipschitz constant $L = ac\rho_-$ into (7.21) in the statement of Theorem 7.2, we obtain the error estimate (7.22).

$x^\star$ is the only solution within a distance ρ_+ of $x^{(0)}$ We claimed that there is only one fixed point of Φ (and solution of (7.1)) in the set $\left\{x \in \mathbb{R}^n \,\middle|\, \left\|x - x^{(0)}\right\| < \rho_+\right\}$. Since we have already proven that there is exactly one fixed point in its subset $\left\{x \in \mathbb{R}^n \,\middle|\, \left\|x - x^{(0)}\right\| \leq \rho_-\right\}$, we must show that there are no fixed points of Φ in $\left\{x \in \mathbb{R}^n \,\middle|\, \rho_- < \left\|x - x^{(0)}\right\| < \rho_+\right\}$. That is, we must show that:

$$
\left(\rho_- < \left\|x - x^{(0)}\right\| < \rho_+\right) \Rightarrow (\Phi(x) \neq x),
$$

or equivalently that:

$$\left(\rho_- < \left\|x - x^{(0)}\right\| < \rho_+\right) \Rightarrow ([J(x^{(0)})]^{-1} g(x) \neq \mathbf{0}).$$

So, let us suppose that $\left(\rho_- < \left\|x - x^{(0)}\right\| < \rho_+\right)$. We prove that this implies that $[J(x^{(0)})]^{-1} g(x) \neq \mathbf{0}$. There are two main steps.

Step 1: From (B.17):

$$\begin{aligned}
&[J(x^{(0)})]^{-1} g(x) \\
&\quad = \quad (x - x^{(0)}) + [J(x^{(0)})]^{-1} g(x^{(0)}) \\
&\qquad\qquad + [J(x^{(0)})]^{-1} \int_{t=0}^{t=1} \left[J(x^{(0)} + t(x - x^{(0)})) - J(x^{(0)})\right] \times (x - x^{(0)})\, dt.
\end{aligned}$$

Therefore, by the triangle inequality (see Exercise A.3) and Lemma A.1:

$$\begin{aligned}
&\left\|[J(x^{(0)})]^{-1} g(x)\right\| \\
&\quad \geq \quad \left\|x - x^{(0)}\right\| - \left\|[J(x^{(0)})]^{-1} g(x^{(0)})\right\| \\
&\qquad\qquad - \left\|[J(x^{(0)})]^{-1}\right\| \times \int_{t=0}^{t=1} \left\|J(x^{(0)} + t(x - x^{(0)})) - J(x^{(0)})\right\| \left\|x - x^{(0)}\right\| dt, \\
&\quad \geq \quad \left\|x - x^{(0)}\right\| - b - a \int_{t=0}^{t=1} \left\|J(x^{(0)} + t(x - x^{(0)})) - J(x^{(0)})\right\| \left\|x - x^{(0)}\right\| dt, \\
&\qquad\qquad \text{by assumption,} \\
&\quad \geq \quad \left\|x - x^{(0)}\right\| - b - a \int_{t=0}^{t=1} ct \left\|x - x^{(0)}\right\|^2 dt, \text{ since:}
\end{aligned}$$

- J is Lipschitz with constant c in the ball of radius $\overline{\rho} \geq \rho_+$ about $x^{(0)}$;
- $x^{(0)} + t(x - x^{(0)})$ is contained in this ball for $0 \leq t \leq 1$; and
- $(x^{(0)} + t(x - x^{(0)})) - x^{(0)} = t(x - x^{(0)})$,

$$= \quad \left\|x - x^{(0)}\right\| - b - ac \left\|x - x^{(0)}\right\|^2 / 2, \text{ on integrating.} \qquad \text{(B.20)}$$

Step 2: We claim that the right-hand side of (B.20) is greater than zero for:

$$\rho_- < \left\|x - x^{(0)}\right\| < \rho_+.$$

Consider the quadratic function:

$$\rho - b - ac(\rho)^2/2. \qquad \text{(B.21)}$$

It has zeros $\rho_- = (1 - \sqrt{1 - 2abc})/(ac)$ and $(1 + \sqrt{1 - 2abc})/(ac)$. Furthermore, the coefficient of $(\rho)^2$ in (B.21) is negative, so (B.21) is positive for ρ in the range $\rho_- < \rho < (1 + \sqrt{1 - 2abc})/(ac)$.

Now let $\rho = \left\| x - x^{(0)} \right\|$. By assumption, $\rho_- < \left\| x - x^{(0)} \right\| < \rho_+$, but $\rho_+ \leq (1 + \sqrt{1 - 2abc})/(ac)$ by definition, so (B.21) is positive for $\rho = \left\| x - x^{(0)} \right\|$ in the range $\rho_- < \rho < \rho_+$. That is:

$$\left\| x - x^{(0)} \right\| - b - ac \left\| x - x^{(0)} \right\|^2 /2 > 0.$$

But, by (B.20), this means that:

$$\left\| [J(x^{(0)})]^{-1} g(x) \right\| \geq \left\| x - x^{(0)} \right\| - b - ac \left\| x - x^{(0)} \right\|^2 /2 > 0.$$

Therefore, there are no fixed points of Φ in:

$$\left\{ x \in \mathbb{R}^n \left| \rho_- < \left\| x - x^{(0)} \right\| < \rho_+ \right. \right\}.$$

□

B.4 Algorithms for linear equality-constrained minimization

Theorem 13.1 First notice that for $x^\star$ to be optimal for the problem it must be feasible, so that $Ax^\star = b$.

Define the function $\tau : \mathbb{R}^{n'} \to \{x \in \mathbb{R}^n | Ax = b\}$ by:

$$\forall \xi \in \mathbb{R}^{n'}, \tau(\xi) = x^\star + Z\xi,$$

which is onto $\{x \in \mathbb{R}^n | Ax = b\}$ by definition of Z. (See Exercise 13.1, Part (ii).) Consider the function $\phi : \mathbb{R}^{n'} \to \mathbb{R}$ defined by:

$$\forall \xi \in \mathbb{R}^{n'}, \phi(\xi) = f(\tau(\xi)).$$

The function ϕ is partially differentiable with continuous partial derivatives since it is the composition of f and τ, which are both partially differentiable with continuous partial derivatives. (See Exercise 13.1, Part (iii).)

By hypothesis, $x^\star \in \operatorname{argmin}_{x \in \mathbb{R}^n} \{f(x) | Ax = b\}$. Therefore, by Theorem 3.5, there exists $\xi^\star \in \operatorname{argmin}_{\xi \in \mathbb{R}^{n'}} \phi(\xi)$ such that $x^\star = \tau(\xi^\star)$.

By Theorem 10.3 applied to the unconstrained problem $\min_{\xi \in \mathbb{R}^{n'}} \phi(\xi)$, we have that $\nabla \phi(\xi^\star) = \mathbf{0}$. But, $\phi(\bullet) = f(\tau(\bullet))$, so:

$$\begin{aligned} \frac{\partial \phi}{\partial \xi}(\xi^\star) &= \frac{\partial f}{\partial x}(\tau(\xi^\star)) \times \frac{\partial \tau}{\partial \xi}(\xi^\star), \text{ by the chain rule [72, section 2.4]}, \\ &= \frac{\partial f}{\partial x}(x^\star) Z, \text{ by definition of } \tau \text{ and by Exercise 13.1, Part (iii)}, \end{aligned}$$

so that $\nabla\phi(\xi^\star) = Z^\dagger \nabla f(x^\star)$. That is, $Z^\dagger \nabla f(x^\star) = \mathbf{0}$. $\Box$

B.5 Algorithms for linear inequality-constrained minimization

Theorem 17.1 ([84, section 14.4].) Consider the *equality*-constrained problem:

$$\min_{x \in \mathbb{R}^n} \{ f(x) | Ax = b, C_\ell x = d_\ell, \forall \ell \in \mathbb{A}(x^\star) \}. \tag{B.22}$$

Problem (B.22) includes all the constraints of Problem (17.1) that were satisfied with equality by $x^\star$. The active inequality constraints from Problem (17.1) have been included as equality constraints in Problem (B.22).

We are going to apply our earlier results for *equality*-constrained problems to Problem (B.22) to prove the theorem. We divide the proof into three parts:

(i) showing that $x^\star$ is a local minimizer of Problem (B.22),
(ii) using the necessary conditions of Problem (B.22) to define $\lambda^\star$ and $\mu^\star$ that satisfy the first four lines of (17.2), and
(iii) proving that $\mu^\star \geq \mathbf{0}$.

$x^\star$ is a local minimizer of Problem (B.22) We prove this by contradiction. Suppose that $x^\star$ is not a local minimum of Problem (B.22). We consider the implications of this supposition.

For any $\ell \notin \mathbb{A}(x^\star)$ we have that $C_\ell x^\star < d_\ell$. By continuity of the continuous function Cx, let $\overline{\epsilon} > 0$ be small enough such that:

$$\forall \ell \notin \mathbb{A}(x^\star), \forall x \text{ such that } \left\| x^\star - x \right\| \leq \overline{\epsilon}, C_\ell x < d_\ell. \tag{B.23}$$

That is, the inequality constraints that are not active at $x^\star$ are also not active at points x that are nearby to $x^\star$.

Let $\epsilon > 0$ be given. By hypothesis, $x^\star$ is not a local minimum of Problem (B.22). Therefore, by (2.27), there exists x^ϵ such that:

$$\begin{aligned} \left\| x^\star - x^\epsilon \right\| &\leq \min\{\epsilon, \overline{\epsilon}\}, \\ &\leq \epsilon, \\ f(x^\epsilon) &< f(x^\star), \\ Ax^\epsilon &= b, \\ \forall \ell \in \mathbb{A}(x^\star), C_\ell x^\epsilon &= d_\ell. \end{aligned}$$

But these, together with (B.23) mean that there is a point x^ϵ that:

- is within ϵ of $x^\star$,
- is feasible for Problem (17.1), and
- has a smaller value of the objective.

Furthermore, such a point x^ϵ exists for any $\epsilon > 0$. Therefore, $x^\star$ is not a local minimum of Problem (17.1). This is a contradiction, so $x^\star$ is in fact a local minimizer of Problem (B.22).

Implications of the necessary conditions of Problem (B.22) Consider Theorem 13.2 applied to Problem (B.22). The objective f is partially differentiable with continuous partial derivatives and $x^\star$ is a local minimizer of Problem (B.22). Therefore, by Theorem 13.2:

$$\exists \lambda^\star \in \mathbb{R}^m, \forall \ell \in \mathbb{A}(x^\star), \exists \mu_\ell^\star \in \mathbb{R} \text{ such that } \nabla f(x^\star) + A^\dagger \lambda^\star + \sum_{\ell \in \mathbb{A}(x^\star)} [C_\ell]^\dagger \mu_\ell^\star = \mathbf{0}. \tag{B.24}$$

We now consider constraints $\ell \in \mathbb{A}(x^\star)$ and constraints $\ell \notin \mathbb{A}(x^\star)$ separately.

By definition, $\forall \ell \in \mathbb{A}(x^\star)$, $C_\ell x^\star = d_\ell$, so that:

$$\forall \ell \in \mathbb{A}(x^\star), \mu_\ell^\star (C_\ell x^\star - d_\ell) = 0. \tag{B.25}$$

Define $\mu_\ell^\star = 0, \forall \ell \notin \mathbb{A}(x^\star)$. Then, trivially, $\forall \ell \notin \mathbb{A}(x^\star)$, $\mu_\ell^\star (C_\ell x^\star - d_\ell) = 0$ and, combining with (B.25), we obtain:

$$\forall \ell = 1, \dots, r, \mu_\ell^\star (C_\ell x^\star - d_\ell) = 0,$$

which is the second line of (17.2). Moreover, $\forall \ell \notin \mathbb{A}(x^\star)$, $[C_\ell]^\dagger \mu_\ell^\star = \mathbf{0}$ so that:

$$\begin{aligned} C^\dagger \mu^\star &= \sum_{\ell \in \mathbb{A}(x^\star)} [C_\ell]^\dagger \mu_\ell^\star + \sum_{\ell \notin \mathbb{A}(x^\star)} [C_\ell]^\dagger \mu_\ell^\star, \\ &= \sum_{\ell \in \mathbb{A}(x^\star)} [C_\ell]^\dagger \mu_\ell^\star. \end{aligned}$$

Therefore, combining with (B.24), we obtain:

$$\exists \lambda^\star \in \mathbb{R}^m, \exists \mu^\star \in \mathbb{R}^r, \text{ such that } \nabla f(x^\star) + A^\dagger \lambda^\star + C^\dagger \mu^\star = \mathbf{0}, \tag{B.26}$$

which is the first line of (17.2).

Non-negativity of $\mu^\star$ By definition, $\forall \ell \notin \mathbb{A}(x^\star)$, $\mu_\ell^\star = 0 \geq 0$. We are left with proving that $\mu_\ell^\star \geq 0, \forall \ell \in \mathbb{A}(x^\star)$. Suppose that this is not true; that is, suppose that $\mu_{\ell'}^\star < 0$ for some $\ell' \in \mathbb{A}(x^\star)$. We construct a step direction Δx and an upper limit on the step size, $\overline{\alpha} > 0$, such that $x^\star + \alpha \Delta x$ is feasible for $0 \leq \alpha \leq \overline{\alpha}$ and f decreases in the direction of Δx away from $x^\star$.

Consider the matrix $\hat{A}$ consisting of all the rows of A together with the rows C_ℓ of C for those $\ell \in \mathbb{A}(x^\star)$. That is, the rows of $\hat{A}$ consist of:

- the m rows of A, and
- those rows of C corresponding to the active constraints.

We assume that $\hat{A}$ has linearly independent rows. (Otherwise, consider a maximal subset of the rows of $\hat{A}$ that are linearly independent and that includes the row corresponding to constraint ℓ'.) Using the analysis in Section 5.8.1.2 we can solve the equation $\hat{A}\Delta x = -\mathbf{I}_{\ell'}$ for Δx, where $\mathbf{I}_{\ell'}$ is a vector that has zeros everywhere except in the position corresponding to inequality constraint ℓ'. We are going to show that $x^\star + \alpha\Delta x$ is feasible for Problem (17.1) for α sufficiently small and positive. To do this, we will, in order, consider feasibility with respect to:

- the equality constraints,
- the inequality constraints that are active at $x^\star$, except for constraint ℓ',
- constraint ℓ', and
- the constraints that are inactive at $x^\star$.

We have that:

$$\begin{aligned}
\forall\alpha \in \mathbb{R}, A(x^\star + \alpha\Delta x) &= Ax^\star + \alpha A\Delta x, \\
&= b + \alpha\mathbf{0}, \\
&\quad \text{by assumption on } x^\star \text{ and construction of } \Delta x, \\
&= b, \\
\forall\alpha \in \mathbb{R}, \forall\ell \in \mathbb{A}(x^\star) \setminus \{\ell'\}, & \\
C_\ell(x^\star + \alpha\Delta x) &= C_\ell x^\star + \alpha C_\ell\Delta x, \\
&= d_\ell + \alpha 0, \\
&\quad \text{by assumption on } x^\star \text{ and construction of } \Delta x, \\
&= d_\ell, \\
&\leq d_\ell, \\
\forall\alpha \geq 0, C_{\ell'}(x^\star + \alpha\Delta x) &= C_{\ell'}x^\star + \alpha C_{\ell'}\Delta x, \\
&= d_{\ell'} + \alpha C_{\ell'}\Delta x, \\
&= d_{\ell'} + \alpha(-1), \text{ by construction of } \Delta x, \\
&< d_{\ell'}.
\end{aligned}$$

By continuity, $\exists\overline{\alpha} > 0$ such that:

$$\begin{aligned}
\forall\ell \notin \mathbb{A}(x^\star), C_\ell(x^\star + \alpha\Delta x) &= C_\ell x^\star + \alpha C_\ell\Delta x, \\
&< d_\ell + \alpha C_\ell\Delta x, \text{ since } \ell \notin \mathbb{A}(x^\star), \\
&\leq d_\ell, \text{ for } 0 \leq \alpha \leq \overline{\alpha}.
\end{aligned}$$

That is, movement in the direction Δx is feasible for step-sizes $0 \leq \alpha \leq \overline{\alpha}$. Moreover,

$$\begin{aligned}
\nabla f(x^\star)^\dagger \Delta x &= -[\lambda^\star]^\dagger A \Delta x - [\mu^\star]^\dagger C \Delta x, \text{ by (B.26)},\\
&= -\mu^\star_{\ell'} C_{\ell'} \Delta x, \text{ by construction of } \Delta x,\\
&= -\mu^\star_{\ell'}(-1), \text{ by construction of } \Delta x,\\
&< 0, \text{ since } \mu^\star_{\ell'} < 0 \text{ by assumption.}
\end{aligned}$$

But this means that f decreases in the direction Δx from $x^\star$ and there are feasible steps in this direction. This contradicts the local optimality of $x^\star$. Therefore, no such ℓ' exists and so $\mu \geq \mathbf{0}$. $\square$

Theorem 17.3 By Item (iv), $x^\star$ is feasible. Consider any other feasible point $x' \in \mathbb{R}^n$. That is, consider x' such that:

$$Ax' = b, Cx' \leq d.$$

We have $Ax' = Ax^\star = b$, so $A(x' - x^\star) = \mathbf{0}$ and:

$$[\lambda^\star]^\dagger A(x' - x^\star) = \mathbf{0}. \tag{B.27}$$

We now consider constraints $\ell \in \mathbb{A}(x^\star)$ and constraints $\ell \notin \mathbb{A}(x^\star)$ separately.

For $\ell \notin \mathbb{A}(x^\star)$, $C_\ell x^\star < d_\ell$ and Item (iii) implies that $\mu^\star_\ell = 0$. Therefore,

$$\forall \ell \notin \mathbb{A}(x^\star), \mu^\star_\ell C_\ell (x' - x^\star) = 0. \tag{B.28}$$

Also, since $C_\ell x' \leq d_\ell$ for all ℓ and since $C_\ell x^\star = d_\ell$ for $\ell \in \mathbb{A}(x^\star)$, we have:

$$\begin{aligned}
\forall \ell \in \mathbb{A}(x^\star), C_\ell(x' - x^\star) &= C_\ell x' - d_\ell,\\
&\leq d_\ell - d_\ell,\\
&= 0.
\end{aligned}$$

Therefore, since $\mu^\star_\ell \geq 0$ for $\ell \in \mathbb{A}(x^\star)$, we have:

$$\forall \ell \in \mathbb{A}(x^\star), \mu^\star_\ell C_\ell (x' - x^\star) \leq 0. \tag{B.29}$$

We have:

$$\begin{aligned}
f(x') &\geq f(x^\star) + \nabla f(x^\star)^\dagger (x' - x^\star), \text{ by Theorem 2.6, noting that:} \\
&\quad f \text{ is partially differentiable with continuous partial derivatives;} \\
&\quad \text{by Item (i) of the hypothesis,} \\
&\quad f \text{ is convex on the convex set } \{x \in \mathbb{R}^n | Ax = b, Cx \leq d\}; \text{ and} \\
&\quad \text{by Item (iv) of the hypothesis and construction,} \\
&\quad x', x^\star \in \{x \in \mathbb{R}^n | Ax = b, Cx \leq d\}, \\
&= f(x^\star) - [A^\dagger \lambda^\star + C^\dagger \mu^\star]^\dagger (x' - x^\star), \\
&\qquad \text{by Item (ii) of the hypothesis,} \\
&= f(x^\star) - [\lambda^\star]^\dagger A(x' - x^\star) - [\mu^\star]^\dagger C(x' - x^\star), \\
&= f(x^\star) - [\mu^\star]^\dagger C(x^\star)(x' - x^\star), \text{ by (B.27),} \\
&= f(x^\star) - \sum_{\ell \in \mathbb{A}(x^\star)} \mu_\ell^\star C_\ell (x' - x^\star) - \sum_{\ell \notin \mathbb{A}(x^\star)} \mu_\ell^\star C_\ell (x' - x^\star), \\
&= f(x^\star) - \sum_{\ell \in \mathbb{A}(x^\star)} \mu_\ell^\star C_\ell (x' - x^\star), \text{ by (B.28),} \\
&\geq f(x^\star), \text{ by (B.29).}
\end{aligned}$$

Therefore $x^\star$ is a global minimizer of f on $\{x \in \mathbb{R}^n | Ax = b, Cx \leq d\}$. □

B.6 Algorithms for non-linear inequality-constrained minimization

Theorem 19.4 By Item (v), $x^\star$ is feasible. Consider any other feasible point $x' \in \mathbb{R}^n$. That is, consider x' such that:

$$Ax' = b, h(x') \leq \mathbf{0}.$$

We have $Ax' = Ax^\star = b$, so $A(x' - x^\star) = \mathbf{0}$ and:

$$[\lambda^\star]^\dagger A(x' - x^\star) = \mathbf{0}. \tag{B.30}$$

We now consider constraints $\ell \in \mathbb{A}(x^\star)$ and constraints $\ell \notin \mathbb{A}(x^\star)$ separately.

For $\ell \notin \mathbb{A}(x^\star)$, $h(x^\star) < 0$ and Item (iv) implies that $\mu_\ell^\star = 0$. Therefore,

$$\forall \ell \notin \mathbb{A}(x^\star), \mu_\ell^\star K_\ell(x^\star)(x' - x^\star) = 0, \tag{B.31}$$

where K_ℓ is the ℓ-th row of K. Also, since $h_\ell(x') \leq 0$ for all ℓ and since $h_\ell(x^\star) = 0$

for $\ell \in \mathbb{A}(x^\star)$, we have:

$$\begin{aligned}\forall \ell \in \mathbb{A}(x^\star), h_\ell(x') - h(x^\star) &= h_\ell(x') - 0,\\ &\leq 0.\end{aligned}$$

We have that:

$$h_\ell(x') \geq h_\ell(x^\star) + K_\ell(x^\star)(x' - x^\star),$$

by Theorem 2.6, noting that h_ℓ is partially differentiable with continuous partial derivatives and is convex by Item (i). Therefore, since $\mu_\ell^\star \geq 0$ for $\ell \in \mathbb{A}(x^\star)$, we have:

$$\forall \ell \in \mathbb{A}(x^\star), \mu_\ell^\star K_\ell(x^\star)(x' - x^\star) \leq 0. \tag{B.32}$$

By Item (i), h is convex so that $\{x \in \mathbb{R}^n | Ax = b, h(x) \leq \mathbf{0}\}$ is a convex set. We have:

$$\begin{aligned}
f(x') &\geq f(x^\star) + \nabla f(x^\star)^\dagger (x' - x^\star), \text{ by Theorem 2.6, noting that:}\\
&\quad f \text{ is partially differentiable with continuous partial derivatives;}\\
&\quad \text{by Item (ii) of the hypothesis,}\\
&\quad f \text{ is convex on the convex set } \{x \in \mathbb{R}^n | Ax = b, h(x) \leq \mathbf{0}\}\text{; and}\\
&\quad \text{by Item (v) of the hypothesis and construction,}\\
&\quad x', x^\star \in \{x \in \mathbb{R}^n | Ax = b, h(x) \leq \mathbf{0}\},\\
&= f(x^\star) - [A^\dagger \lambda^\star + K(x^\star)^\dagger \mu^\star]^\dagger (x' - x^\star),\\
&\qquad\qquad \text{by Item (iii) of the hypothesis,}\\
&= f(x^\star) - [\lambda^\star]^\dagger A(x' - x^\star) - [\mu^\star]^\dagger K(x^\star)(x' - x^\star),\\
&= f(x^\star) - [\mu^\star]^\dagger K(x^\star)(x' - x^\star), \text{ by (B.30)},\\
&= f(x^\star) - \sum_{\ell \in \mathbb{A}(x^\star)} \mu_\ell^\star K_\ell(x^\star)(x' - x^\star) - \sum_{\ell \notin \mathbb{A}(x^\star)} \mu_\ell^\star K_\ell(x^\star)(x' - x^\star),\\
&= f(x^\star) - \sum_{\ell \in \mathbb{A}(x^\star)} \mu_\ell^\star K_\ell(x^\star)(x' - x^\star), \text{ by (B.31)},\\
&\geq f(x^\star), \text{ by (B.32)}.
\end{aligned}$$

Therefore $x^\star$ is a global minimizer of f on $\{x \in \mathbb{R}^n | Ax = b, h(x) \leq \mathbf{0}\}$. □