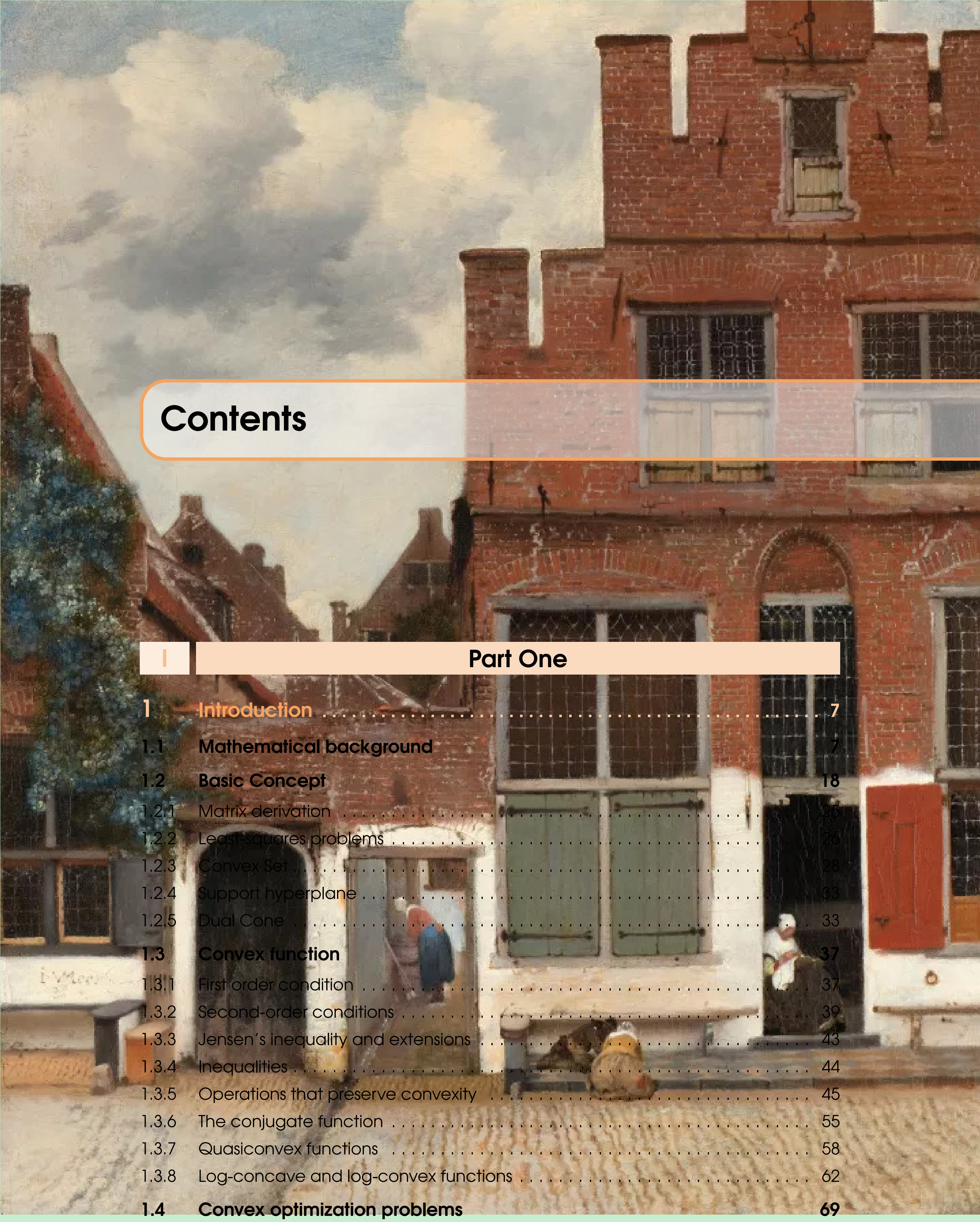
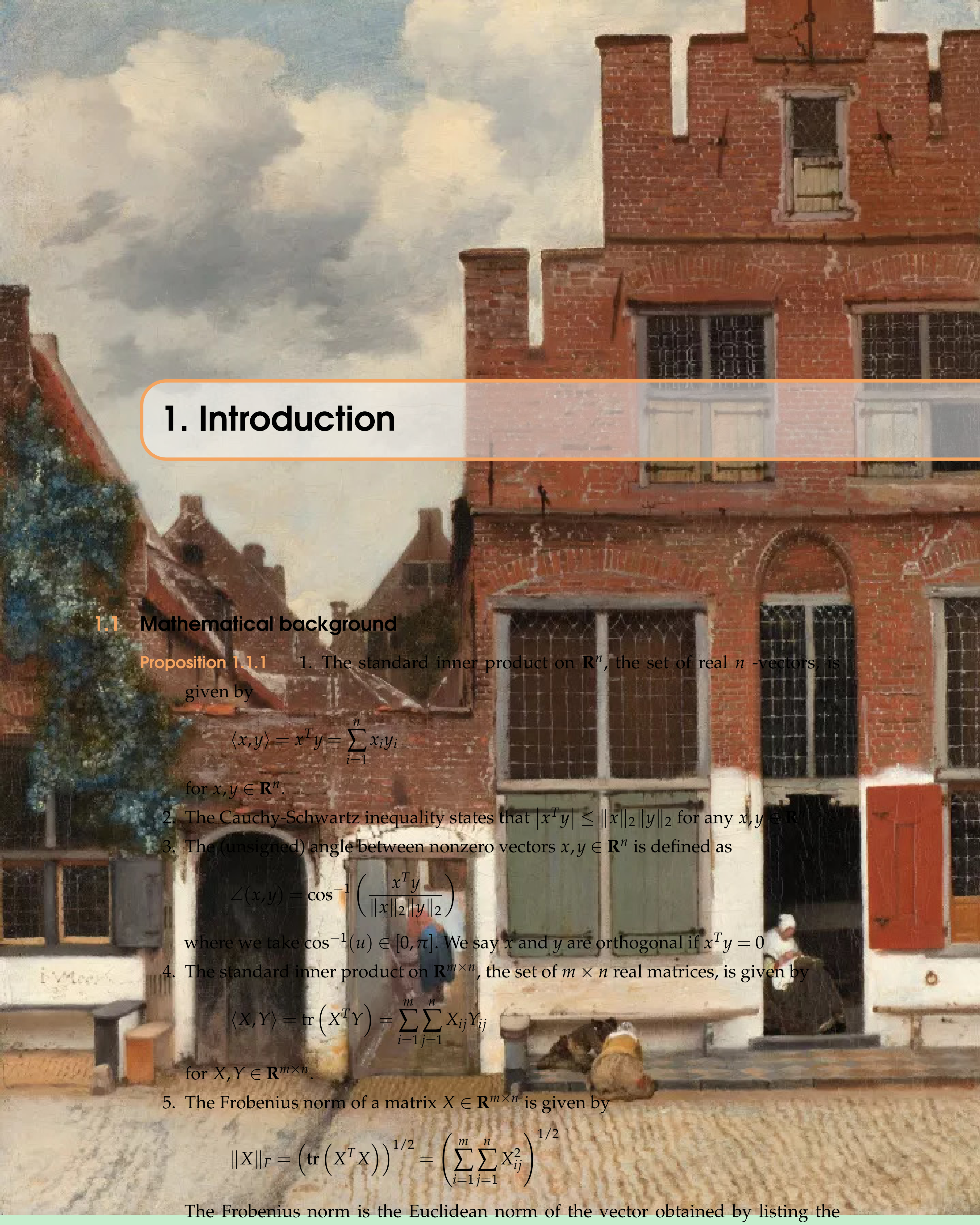
Convex Optimization

Wen Xin



Contents

Part One



1. Introduction

1.1 Mathematical background

Proposition 1.1.1
1. The standard inner product on $\mathbf{R}^n$, the set of real n -vectors, is given by

$$\langle x, y \rangle = x^T y = \sum_{i=1}^{n} x_i y_i$$

for $x, y \in \mathbf{R}^n$.

2. The Cauchy-Schwartz inequality states that $|x^T y| \leq \|x\|_2 \|y\|_2$ for any $x, y \in \mathbf{R}^n$.
3. The (unsigned) angle between nonzero vectors $x, y \in \mathbf{R}^n$ is defined as

$$\angle(x, y) = \cos^{-1}\left(\frac{x^T y}{\|x\|_2 \|y\|_2}\right)$$

where we take $\cos^{-1}(u) \in [0, \pi]$. We say x and y are orthogonal if $x^T y = 0$

4. The standard inner product on $\mathbf{R}^{m \times n}$, the set of $m \times n$ real matrices, is given by

$$\langle X, Y \rangle = \operatorname{tr}\left(X^T Y\right) = \sum_{i=1}^{m} \sum_{j=1}^{n} X_{ij} Y_{ij}$$

for $X, Y \in \mathbf{R}^{m \times n}$.

5. The Frobenius norm of a matrix $X \in \mathbf{R}^{m \times n}$ is given by

$$\|X\|_F = \left(\operatorname{tr}\left(X^T X\right)\right)^{1/2} = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} X_{ij}^2\right)^{1/2}$$

The Frobenius norm is the Euclidean norm of the vector obtained by listing the coefficients of the matrix.

6. The standard inner product on $\mathbf{S}^n$, the set of symmetric $n \times n$ matrices, is given by

$$\langle X, Y\rangle = \operatorname{tr}(XY) = \sum_{i=1}^{n}\sum_{j=1}^{n} X_{ij}Y_{ij} = \sum_{i=1}^{n} X_{ii}Y_{ii} + 2\sum_{i<j} X_{ij}Y_{ij}$$

7. **Quadratic Norms**: For $P \in \mathbf{S}_{++}^n$, we define the P -quadratic norm as

$$\|x\|_P = \left(x^T P x\right)^{1/2} = \left\|P^{1/2}x\right\|_2$$

The unit ball of a quadratic norm is an ellipsoid (and conversely, if the unit ball of a norm is an ellipsoid, the norm is a quadratic norm).

8. If $\|\cdot\|$ is any norm on $\mathbf{R}^n$, then there exists a quadratic norm $\|\cdot\|_P$ for which

$$\|x\|_P \le \|x\| \le \sqrt{n}\|x\|_P$$

holds for all x. In other words, any norm on $\mathbf{R}^n$ can be uniformly approximated, within a factor of $\sqrt{n}$, by a quadratic norm.

9. Suppose $\|\cdot\|_{\mathrm{a}}$ and $\|\cdot\|_{\mathrm{b}}$ are norms on $\mathbf{R}^m$ and $\mathbf{R}^n$, respectively. We define the operator norm of $X \in \mathbf{R}^{m\times n}$, induced by the norms $\|\cdot\|_{\mathrm{a}}$ and $\|\cdot\|_{\mathrm{b}}$, as

$$\|X\|_{\mathrm{a,b}} = \sup\left\{\|Xu\|_{\mathrm{a}} \mid \|u\|_{\mathrm{b}} \le 1\right\}$$

10. When $\|\cdot\|_{\mathrm{a}}$ and $\|\cdot\|_{\mathrm{b}}$ are both Euclidean norms, the operator norm of X is its maximum singular value, and is denoted $\|X\|_2$:

$$\|X\|_2 = \sigma_{\max}(X) = \left(\lambda_{\max}\left(X^T X\right)\right)^{1/2}$$

(This agrees with the Euclidean norm on $\mathbf{R}^m$, when $X \in \mathbf{R}^{m\times 1}$, so there is no clash of notation.) This norm is also called the **spectral norm** or ℓ_2 -norm of X.

11. The norm induced by the ℓ_∞ -norm on $\mathbf{R}^m$ and $\mathbf{R}^n$, denoted $\|X\|_\infty$, is the max-row-sum norm

$$\|X\|_\infty = \sup\left\{\|Xu\|_\infty \mid \|u\|_\infty \le 1\right\} = \max_{i=1,\ldots,m}\sum_{j=1}^{n}\left|X_{ij}\right|$$

12. The norm induced by the ℓ_1 -norm on $\mathbf{R}^m$ and $\mathbf{R}^n$, denoted $\|X\|_1$, is the max column-sum norm,

$$\|X\|_1 = \max_{j=1,\ldots,n}\sum_{i=1}^{m}\left|X_{ij}\right|$$

13. Let $\|\cdot\|$ be a norm on $\mathbf{R}^n$. The associated **dual norm**, denoted $\|\cdot\|_*$, is defined as

$$\|z\|_* = \sup\left\{z^T x \mid \|x\| \le 1\right\}$$

(This can be shown to be a norm.) The dual norm can be interpreted as the operator norm of z^T, interpreted as a $1 \times n$ matrix, with the norm $\|\cdot\|$ on $\mathbf{R}^n$, and the absolute value on $\mathbf{R}$:

$$\|z\|_* = \sup\left\{\left|z^T x\right| \mid \|x\| \le 1\right\}$$

14. From the definition of dual norm we have the inequality

$$z^T x \leq \|x\| \|z\|_*$$

which holds for all x and z. This inequality is tight, in the following sense: for any x there is a z for which the inequality holds with equality. (Similarly, for any z there is an x that gives equality.)

15. The dual of the dual norm is the original norm: we have $\|x\|_{**} = \|x\|$ for all x. (This need not hold in infinite-dimensional vector spaces.)

16. The dual of the Euclidean norm is the Euclidean norm, since

$$\sup \left\{ z^T x \mid \|x\|_2 \leq 1 \right\} = \|z\|_2$$

(This follows from the Cauchy-Schwarz inequality; for nonzero z, the value of x that maximizes $z^T x$ over $\|x\|_2 \leq 1$ is $z / \|z\|_2$.)

17. The dual of the ℓ_∞ -norm is the ℓ_1 -norm:

$$\sup \left\{ z^T x \mid \|x\|_\infty \leq 1 \right\} = \sum_{i=1}^{n} |z_i| = \|z\|_1$$

and the dual of the ℓ_1 -norm is the ℓ_∞ -norm.

18. More generally, the dual of the ℓ_p -norm is the ℓ_q -norm, where q satisfies $1/p + 1/q = 1$, i.e., $q = p/(p-1)$

19. As another example, consider the ℓ_2- or spectral norm on $\mathbf{R}^{m \times n}$. The associated dual norm is

$$\|Z\|_{2*} = \sup \left\{ \operatorname{tr} \left(Z^T X \right) \mid \|X\|_2 \leq 1 \right\}$$

which turns out to be the sum of the singular values,

$$\|Z\|_{2*} = \sigma_1(Z) + \cdots + \sigma_r(Z) = \operatorname{tr} \left(Z^T Z \right)^{1/2}$$

where $r = \operatorname{rank} Z$. This norm is sometimes called the **nuclear norm**.

20. The **closure** of a set C is defined as

$$\mathbf{cl} C = \mathbf{R}^n \backslash \operatorname{int} (\mathbf{R}^n \backslash C)$$

i.e., the complement of the interior of the complement of C. A point x is in the closure of C if for every $\epsilon > 0$, there is a $y \in C$ with $\|x - y\|_2 \leq \epsilon$. The closure of C is the set of all limit points of convergent sequences in C.

21. The boundary of the set C is defined as

$$\mathbf{bd} C = \mathbf{cl} C \backslash \operatorname{int} C$$

22. A boundary point x (i.e., a point $x \in \operatorname{bd} C$) satisfies the following property: For all $\epsilon > 0$, there exists $y \in C$ and $z \notin C$ with

$$\|y - x\|_2 \le \epsilon, \quad \|z - x\|_2 \le \epsilon$$

i.e., there exist arbitrarily close points in C, and also arbitrarily close points not in C.

23. We can characterize closed and open sets in terms of the boundary operation: C is closed if it contains its boundary, *i.e.*, $\operatorname{bd} C \subseteq C$. It is open if it contains no boundary points, i.e., $C \cap \operatorname{bd} C = \emptyset$.

24. A function $f : \mathbf{R}^n \to \mathbf{R}$ is said to be **closed** if, for each $\alpha \in \mathbf{R}$, the sublevel set

$$\{x \in \operatorname{dom} f \mid f(x) \le \alpha\}$$

is closed. This is equivalent to the condition that the epigraph of f,

$$\operatorname{epi} f = \left\{(x,t) \in \mathbf{R}^{n+1} \mid x \in \operatorname{dom} f, f(x) \le t\right\}$$

is closed. (This definition is general, but is usually only applied to convex functions.)

25. If $f : \mathbf{R}^n \to \mathbf{R}$ is continuous, and $\operatorname{dom} f$ is closed, then f is closed. If $f : \mathbf{R}^n \to \mathbf{R}$ is continuous, with $\operatorname{dom} f$ open, then f is closed if and only if f converges to ∞ along every sequence converging to a boundary point of $\operatorname{dom} f$. In other words, if $\lim_{i\to\infty} x_i = x \in \operatorname{bd} \operatorname{dom} f$, with $x_i \in \operatorname{dom} f$, we have $\lim_{i\to\infty} f(x_i) = \infty$

26. **Derivative and gradient**. Suppose $f : \mathbf{R}^n \to \mathbf{R}^m$ and $x \in \operatorname{int} \operatorname{dom} f$. The function f is **differentiable** at x if there exists a matrix $Df(x) \in \mathbf{R}^{m\times n}$ that satisfies

$$\lim_{z\in \operatorname{dom} f, z\neq x, z\to x} \frac{\|f(z) - f(x) - Df(x)(z - x)\|_2}{\|z - x\|_2} = 0$$

in which case we refer to $Df(x)$ as the derivative (or Jacobian) of f at x. (There can be at most one matrix that satisfies previous formula.)

27. The function f is differentiable if $\operatorname{dom} f$ is open, and it is differentiable at every point in its domain.

28. The affine function of z given by

$$f(x) + Df(x)(z - x)$$

is called the **first-order approximation** of f at (or near) x. Evidently this function agrees with f at $z = x$; when z is close to x, this affine function is very close to f. The derivative can be found by deriving the first-order approximation of the function f at x (*i.e.*, the matrix $Df(x)$), or from partial derivatives:

$$Df(x)_{ij} = \frac{\partial f_i(x)}{\partial x_j}, \quad i = 1,\ldots,m, \quad j = 1,\ldots,n$$

29. Example: As a more interesting example, we consider the function $f : \mathbf{S}^n \to \mathbf{R}$, given by

$$f(X) = \log\det X, \quad \operatorname{dom} f = \mathbf{S}^n_{++}$$

We will directly find the first-order approximation of f at $X \in \mathbf{S}^n_{++}$. Let $Z \in \mathbf{S}^n_{++}$ be close to X, and let $\Delta X = Z - X$(which is assumed to be small). We have

$$\begin{aligned}\log\det Z &= \log\det(X + \Delta X)\\ &= \log\det\left(X^{1/2}\left(I + X^{-1/2}\Delta X X^{-1/2}\right)X^{1/2}\right)\\ &= \log\det X + \log\det\left(I + X^{-1/2}\Delta X X^{-1/2}\right)\\ &= \log\det X + \sum_{i=1}^{n}\log\left(1 + \lambda_i\right)\end{aligned}$$

where λ_i is the i th eigenvalue of $X^{-1/2}\Delta X X^{-1/2}$. Now we use the fact that ΔX is small, which implies λ_i are small, so to first order we have $\log\left(1 + \lambda_i\right) \approx \lambda_i$. Using this first-order approximation in the expression above, we get

$$\begin{aligned}\log\det Z &\approx \log\det X + \sum_{i=1}^{n}\lambda_i\\ &= \log\det X + \operatorname{tr}\left(X^{-1/2}\Delta X X^{-1/2}\right)\\ &= \log\det X + \operatorname{tr}\left(X^{-1}\Delta X\right)\\ &= \log\det X + \operatorname{tr}\left(X^{-1}(Z - X)\right)\end{aligned}$$

where we have used the fact that the sum of the eigenvalues is the trace, and the property $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ Thus, the first-order approximation of f at X is the affine function of Z given by

$$f(Z) \approx f(X) + \operatorname{tr}\left(X^{-1}(Z - X)\right)$$

Noting that the second term on the righthand side is the standard inner product of X^{-1} and $Z - X$, we can identify X^{-1} as the gradient of f at X. Thus, we can write the simple formula

$$\nabla f(X) = X^{-1}$$

This result should not be surprising, since the derivative of $\log x$, on $\mathbf{R}_{++}$, is $1/x$

30. **Gradient**: When f is real-valued (*i.e.*, $f : \mathbf{R}^n \to \mathbf{R}$) the derivative $Df(x)$ is a $1 \times n$ matrix, i.e., it is a row vector. Its transpose is called the gradient of the function:

$$\nabla f(x) = Df(x)^T$$

which is a (column) vector, i.e., in $\mathbf{R}^n$. Its components are the partial derivatives of f:

$$\nabla f(x)_i = \frac{\partial f(x)}{\partial x_i}, \quad i = 1, \ldots, n$$

The first-order approximation of f at a point $x \in \operatorname{int} \operatorname{dom} f$ can be expressed as (the affine function of z)

$$f(x) + \nabla f(x)^T (z - x)$$

31. **Chain rule**: Suppose $f : \mathbf{R}^n \to \mathbf{R}^m$ is differentiable at $x \in \operatorname{int} \operatorname{dom} f$ and $g : \mathbf{R}^m \to \mathbf{R}^p$ is differentiable at $f(x) \in \operatorname{int} \operatorname{dom} g$. Define the composition $h : \mathbf{R}^n \to \mathbf{R}^p$ by $h(z) = g(f(z))$. Then h is differentiable at x, with derivative

$$Dh(x) = Dg(f(x))Df(x)$$

32. As an example, suppose $f : \mathbf{R}^n \to \mathbf{R}, g : \mathbf{R} \to \mathbf{R}$, and $h(x) = g(f(x))$. Taking the transpose of $Dh(x) = Dg(f(x))Df(x)$ yields

$$\nabla h(x) = g'(f(x)) \nabla f(x)$$

33. **Composition with affine function**: Suppose $f : \mathbf{R}^n \to \mathbf{R}^m$ is differentiable, $A \in \mathbf{R}^{n \times p}$, and $b \in \mathbf{R}^n$. Define $g : \mathbf{R}^p \to \mathbf{R}^m$ as $g(x) = f(Ax + b)$, with $\operatorname{dom} g = \{x \mid Ax + b \in \operatorname{dom} f\}$. The derivative of g is, by the chain rule ,$Dg(x) = Df(Ax + b)A$. When f is real-valued (i.e., $m = 1$), we obtain the formula for the gradient of a composition of a function with an affine function,

$$\nabla g(x) = A^T \nabla f(Ax + b)$$

34. For example, suppose that $f : \mathbf{R}^n \to \mathbf{R}, x, v \in \mathbf{R}^n$, and we define the function $\tilde{f} : \mathbf{R} \to \mathbf{R}$ by $\tilde{f}(t) = f(x + tv)$. (Roughly speaking, $\tilde{f}$ is f, restricted to the line $\{x + tv \mid t \in \mathbf{R}\}$.) Then we have

$$D\tilde{f}(t) = \tilde{f}'(t) = \nabla f(x + tv)^T v$$

(The scalar $\tilde{f}'(0)$ is the directional derivative of f, at x, in the direction v.)

35. **Second derivative**:$f : \mathbf{R}^n \to$ R. The second derivative or **Hessian matrix** of f at $x \in \operatorname{int} \operatorname{dom} f$, denoted $\nabla^2 f(x)$, is given by

$$\nabla^2 f(x)_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}, \quad i = 1, \ldots n, \quad j = 1, \ldots, n$$

provided f is twice differentiable at x, where the partial derivatives are evaluated at x. The second-order approximation of f, at or near x, is the quadratic function of z defined by

$$\widehat{f}(z) = f(x) + \nabla f(x)^T (z - x) + (1/2)(z - x)^T \nabla^2 f(x)(z - x)$$

This second-order approximation satisfies

$$\lim_{z\in \mathrm{dom}\, f, z\neq x, z\to x} \frac{|f(z)-\widehat{f}(z)|}{\|z-x\|_2^2} = 0$$

Not surprisingly, the second derivative can be interpreted as the derivative of the first derivative. If f is differentiable, the gradient mapping is the function $\nabla f : \mathbf{R}^n \to \mathbf{R}^n$, with $\mathrm{dom}\, \nabla f = \mathrm{dom}\, f$, with value $\nabla f(x)$ at x. The derivative of this mapping is

$$D\nabla f(x) = \nabla^2 f(x)$$

36. **Composition with scalar function**: Suppose $f : \mathbf{R}^n \to \mathbf{R}, g : \mathbf{R} \to \mathbf{R}$, and $h(x) = g(f(x))$. Simply working out the partial derivatives yields

$$\nabla^2 h(x) = g'(f(x))\nabla^2 f(x) + g''(f(x))\nabla f(x)\nabla f(x)^T$$

37. **Composition with affine function**: Suppose $f : \mathbf{R}^n \to \mathbf{R}, A \in \mathbf{R}^{n\times m}$, and $b \in \mathbf{R}^n$. Define $g : \mathbf{R}^m \to \mathbf{R}$ by $g(x) = f(Ax+b)$. Then we have

$$\nabla^2 g(x) = A^T \nabla^2 f(Ax+b)A$$

As an example, consider the restriction of a real-valued function f to a line, *i.e.* the function $\tilde{f}(t) = f(x+tv)$, where x and v are fixed. Then we have

$$\nabla^2 \tilde{f}(t) = \tilde{f}''(t) = v^T \nabla^2 f(x+tv)v$$

38. **Orthogonal decomposition induced by** A: If $\mathcal{V}$ is a subspace of $\mathbf{R}^n$, its orthogonal complement, denoted $\mathcal{V}^\perp$, is defined as

$$\mathcal{V}^\perp = \left\{ x \mid z^T x = 0 \text{ for all } z \in \mathcal{V} \right\}$$

(As one would expect of a complement, we have $\mathcal{V}^{\perp\perp} = \mathcal{V}$.)

39. A basic result of linear algebra is that, for any $A \in \mathbf{R}^{m\times n}$, we have

$$\mathcal{N}(A) = \mathcal{R}\left(A^T\right)^\perp$$

(Applying the result to A^T we also have $\mathcal{R}(A) = \mathcal{N}\left(A^T\right)^\perp$.) This result is often stated as

$$\mathcal{N}(A) \stackrel{\perp}{\oplus} \mathcal{R}\left(A^T\right) = \mathbf{R}^n$$

Here the symbol $\stackrel{\perp}{\oplus}$ refers to orthogonal direct sum, i.e., the sum of two subspaces that are orthogonal. The decomposition of $\mathbf{R}^n$ is called **the orthogonal decomposition induced by** A.

40. **Symmetric eigenvalue decomposition**: Suppose $A \in \mathbf{S}^n$, i.e., A is a real symmetric $n \times n$ matrix. Then A can be factored *as*

$$A = Q\Lambda Q^T$$

where $Q \in \mathbf{R}^{n\times n}$ is orthogonal, i.e., satisfies $Q^TQ = I$, and $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.

41. The (real) numbers λ_i are the eigenvalues of A, and are the roots of the characteristic polynomial $\det(sI - A)$. The columns of Q form an orthonormal set of eigenvectors of A. The factorization is called the **spectral decomposition or (symmetric) eigenvalue decomposition of** A**.**
42. We order the eigenvalues as $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. We use the notation $\lambda_i(A)$ to refer to the i th largest eigenvalue of $A \in S$. We usually write the largest or maximum eigenvalue as $\lambda_1(A) = \lambda_{\max}(A)$, and the least or minimum eigenvalue as $\lambda_n(A) = \lambda_{\min}(A)$.
43. The determinant and trace can be expressed in terms of the eigenvalues,

$$\det A = \prod_{i=1}^{n} \lambda_i, \quad \operatorname{tr} A = \sum_{i=1}^{n} \lambda_i$$

as can the spectral and Frobenius norms,

$$\|A\|_2 = \max_{i=1,\ldots,n} |\lambda_i| = \max\{\lambda_1, -\lambda_n\}, \quad \|A\|_F = \left(\sum_{i=1}^{n} \lambda_i^2\right)^{1/2}$$

44. **Definiteness and matrix inequalities**: The largest and smallest eigenvalues satisfy

$$\lambda_{\max}(A) = \sup_{x\neq 0} \frac{x^TAx}{x^Tx}, \quad \lambda_{\min}(A) = \inf_{x\neq 0} \frac{x^TAx}{x^Tx}$$

In particular, for any x, we have

$$\lambda_{\min}(A)x^Tx \leq x^TAx \leq \lambda_{\max}(A)x^Tx$$

with both inequalities tight for (different) choices of x.

45. A matrix $A \in \mathbf{S}^n$ is called **positive definit**e if for all $x \neq 0, x^TAx > 0$. We denote this as $A \succ 0$. By the inequality above, we see that $A \succ 0$ if and only all its eigenvalues are positive, i.e.,$\lambda_{\min}(A) > 0$.
46. If $-A$ is positive definite, we say A is **negative definite**, which we write as $A \prec 0$.
47. We use $\mathbf{S}^n_{++}$ to denote the set of positive definite matrices in $\mathbf{S}^n$.
48. If A satisfies $x^TAx \geq 0$ for all x, we say that A is positive semidefinite or nonnegative definite.
49. If $-A$ is nonnegative definite, i.e., if $x^TAx \leq 0$ for all x, we say that A is negative semidefinite or nonpositive definite.
50. We use $\mathbf{S}^n_{+}$ to denote the set of nonnegative definite matrices in $\mathbf{S}^n$.

51. For $A, B \in \mathbf{S}^n$, we use $A \prec B$ to mean $B - A \succ 0$, and so on. These inequalities are called **matrix inequalities**, or generalized inequalities associated with the positive semidefinite cone.
52. **Symmetric squareroot**: Let $A \in \mathbf{S}^n_+$, with eigenvalue decomposition $A = Q \operatorname{diag}(\lambda_1, \ldots, \lambda_n) Q^T$. We define the (symmetric) squareroot of A as

$$A^{1/2} = Q \operatorname{diag}\left(\lambda_1^{1/2}, \ldots, \lambda_n^{1/2}\right) Q^T$$

The squareroot $A^{1/2}$ is the unique symmetric positive semidefinite solution of the equation $X^2 = A$
53. **Generalized eigenvalue decomposition**: The generalized eigenvalues of a pair of symmetric matrices $(A, B) \in \mathbf{S}^n \times \mathbf{S}^n$ are defined as the roots of the polynomial $\det(sB - A)$. We are usually interested in matrix pairs with $B \in \mathbf{S}^n_{++}$. In this case the generalized eigenvalues are also the eigenvalues of $B^{-1/2} A B^{-1/2}$ (which are real). As with the standard eigenvalue decomposition, we order the generalized eigenvalues in nonincreasing order, as $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, and denote the maximum generalized eigenvalue by $\lambda_{\max}(A, B)$. When $B \in \mathbf{S}^n_{++}$, the pair of matrices can be factored as

$$A = V \Lambda V^T, \quad B = V V^T \tag{1.1}$$

where $V \in \mathbf{R}^{n \times n}$ is nonsingular, and $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, where λ_i are the generalized eigenvalues of the pair (A, B). The decomposition is called the **generalized eigenvalue decomposition.**
54. The generalized eigenvalue decomposition is related to the standard eigenvalue decomposition of the matrix $B^{-1/2} A B^{-1/2}$. If $Q \Lambda Q^T$ is the eigenvalue decomposition of $B^{-1/2} A B^{-1/2}$, then 1.1 holds with $V = B^{1/2} Q$.
55. **Singular value decomposition**: Suppose $A \in \mathbf{R}^{m \times n}$ with $\operatorname{rank} A = r$. Then A can be factored as

$$A = U \Sigma V^T \tag{1.2}$$

where $U \in \mathbf{R}^{m \times r}$ satisfies $U^T U = I$, $V \in \mathbf{R}^{n \times r}$ satisfies $V^T V = I$, and $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$, with

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$$

The factorization (1.2) is called the singular value decomposition (SVD) of A. The columns of U are called left singular vectors of A, the columns of V are right singular vectors, and the numbers σ_i are the singular values. The singular value decomposition can be written

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

where $u_i \in \mathbf{R}^m$ are the left singular vectors, and $v_i \in \mathbf{R}^n$ are the right singular vectors.

56. The singular value decomposition of a matrix A is closely related to the eigenvalue decomposition of the (symmetric, nonnegative definite) matrix $A^T A$. Using 1.2 we can write

$$A^T A = V\Sigma^2 V^T = \begin{bmatrix} V & \tilde{V} \end{bmatrix} \begin{bmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V & \tilde{V} \end{bmatrix}^T$$

where $\tilde{V}$ is any matrix for which $[V \quad \tilde{V}]$ is orthogonal. The righthand expression is the eigenvalue decomposition of $A^T A$, so we conclude that its nonzero eigenvalues are the singular values of A squared, and the associated eigenvectors of $A^T A$ are the right singular vectors of A.

57. A similar analysis of AA^T shows that its nonzero eigenvalues are also the squares of the singular values of A, and the associated eigenvectors are the left singular vectors of A.

58. The first or largest singular value is also written as $\sigma_{\max}(A)$. It can be expressed *as*

$$\sigma_{\max}(A) = \sup_{x,y\neq 0} \frac{x^T A y}{\|x\|_2 \|y\|_2} = \sup_{y\neq 0} \frac{\|Ay\|_2}{\|y\|_2}$$

The righthand expression shows that the maximum singular value is the ℓ_2 operator norm of A.

59. The minimum singular value of $A \in \mathbf{R}^{m\times n}$ is given by

$$\sigma_{\min}(A) = \begin{cases} \sigma_r(A) & r = \min\{m,n\} \\ 0 & r < \min\{m,n\} \end{cases}$$

which is positive if and only if A is full rank.

60. The singular values of a symmetric matrix are the absolute values of its nonzero eigenvalues, sorted into descending order. The singular values of a symmetric positive semidefinite matrix are the same as its nonzero eigenvalues.

61. The condition number of a nonsingular $A \in \mathbf{R}^{n\times n}$, denoted cond (A) or $\kappa(A)$ is defined as

$$\mathrm{cond}(A) = \|A\|_2 \left\|A^{-1}\right\|_2 = \sigma_{\max}(A)/\sigma_{\min}(A)$$

62. **Pseudo-inverse**: Let $A = U\Sigma V^T$ be the singular value decomposition of $A \in \mathbf{R}^{m\times n}$, with rank $A = r$. We define the **pseudo-inverse or Moore-Penrose inverse** of A as

$$A^\dagger = V\Sigma^{-1}U^T \in \mathbf{R}^{n\times m}$$

Alternative expressions are

$$A^\dagger = \lim_{\epsilon\to 0} \left(A^T A + \epsilon I\right)^{-1} A^T = \lim_{\epsilon\to 0} A^T \left(AA^T + \epsilon I\right)^{-1}$$

where the limits are taken with $\epsilon > 0$, which ensures that the inverses in the expressions exist.

63. If $\operatorname{rank} A = n$, then $A^\dagger = \left(A^TA\right)^{-1}A^T$. If $\operatorname{rank} A = m$, then $A^\dagger = A^T\left(AA^T\right)^{-1}$. If A is square and nonsingular, then $A^\dagger = A^{-1}$
The pseudo-inverse comes up in problems involving least-squares, minimum norm, quadratic minimization, and (Euclidean) projection. For example, $A^\dagger b$ is a solution of the least-squares problem
$$\text{minimize } \|Ax - b\|_2^2$$
in general. When the solution is not unique, $A^\dagger b$ gives the solution with minimum (Euclidean) norm.
64. As another example, the matrix $AA^\dagger = UU^T$ gives (Euclidean) projection on $\mathcal{R}(A)$. The matrix $A^\dagger A = VV^T$ gives (Euclidean) projection on $\mathcal{R}\left(A^T\right)$
65. The optimal value $p^\star$ of the (general, nonconvex) quadratic optimization problem
$$\text{minimize} \quad (1/2)x^TPx + q^Tx + r$$
where $P \in \mathbf{S}^n$, can be expressed as
$$p^\star = \begin{cases} -(1/2)q^TP^\dagger q + r & P \succeq 0, \quad q \in \mathcal{R}(P) \\ -\infty & \text{otherwise} \end{cases}$$
(This generalizes the expression $p^\star = -(1/2)q^TP^{-1}q + r$, valid for $P \succ 0$.)
66. **Schur complement**: Consider a matrix $X \in \mathbf{S}^n$ partitioned as
$$X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$$
where $A \in \mathbf{S}^k$. If $\det A \neq 0$, the matrix
$$S = C - B^TA^{-1}B$$
is called the **Schur complement** of A in X. Schur complements arise in several contexts, and appear in many important formulas and theorems. For example, we have
$$\det X = \det A \det S$$
67. **Inverse of block matrix**: The Schur complement comes up in solving linear equations, by eliminating one block of variables. We start with
$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix}$$
and assume that $\det A \neq 0$. If we eliminate x from the top block equation and substitute it into the bottom block equation, we obtain $v = B^TA^{-1}u + Sy$, so
$$y = S^{-1}\left(v - B^TA^{-1}u\right)$$

Substituting this into the first equation yields

$$x = \left(A^{-1} + A^{-1}BS^{-1}B^TA^{-1}\right)u - A^{-1}BS^{-1}v$$

We can express these two equations as a formula for the inverse of a block matrix:

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} + A^{-1}BS^{-1}B^TA^{-1} & -A^{-1}BS^{-1} \\ -S^{-1}B^TA^{-1} & S^{-1} \end{bmatrix}$$

In particular, we see that the Schur complement is the inverse of the 2,2 block entry of the inverse of X.

68. **Minimization and definiteness**: The Schur complement arises when you minimize a quadratic form over some of the variables. Suppose $A \succ 0$, and consider the minimization problem

$$\text{minimize} \quad u^TAu + 2v^TB^Tu + v^TCv$$

with variable u. The solution is $u = -A^{-1}Bv$, and the optimal value is

$$\inf_u \begin{bmatrix} u \\ v \end{bmatrix}^T \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = v^TSv$$

From this we can derive the following characterizations of positive definiteness or semidefiniteness of the block matrix X :

(a) $\boldsymbol{X} \succ 0$ if and only if $A \succ 0$ and $S \succ 0$

(b) If $A \succ 0$, then $X \succeq 0$ if and only if $S \succeq 0$.

69. **Schur complement with singular** A: Some Schur complement results have generalizations to the case when A is singular, although the details are more complicated. As an example, if $A \succeq 0$ and $Bv \in \mathcal{R}(A)$, then the quadratic minimization problem (A.13) (with variable u) is solvable, and has optimal value

$$v^T\left(C - B^TA^{\dagger}B\right)v$$

where $A^{\dagger}$ is the pseudo-inverse of A. The problem is unbounded if $Bv \notin \mathcal{R}(A)$ or if $A \nsucceq 0$. The range condition $Bv \in \mathcal{R}(A)$ can also be expressed as $\left(I - AA^{\dagger}\right)Bv = 0$ so we have the following characterization of positive semidefiniteness of the block matrix X :

$$X \succeq 0 \quad \Longleftrightarrow \quad A \succeq 0, \quad \left(I - AA^{\dagger}\right)B = 0, \quad C - B^TA^{\dagger}B \succeq 0$$

Here the matrix $C - B^TA^{\dagger}B$ serves as a generalization of the Schur complement, when A is singular.

1.2 Basic Concept

Definition 1.2.1 — Basic Concepts.

1. A problem is **sparse** if each constraint function depends on only a small number of the variables.
2. A set $C \subseteq \mathbf{R}^n$ is **affine** if the line through any two distinct points in C lies in C. Or if C is an **affine** set, $x_1, \ldots, x_k \in C$, and $\theta_1 + \cdots + \theta_k = 1$, then the point $\theta_1 x_1 + \cdots + \theta_k x_k$ also belongs to C
3. A set C is **convex** if the line segment between any two points in C lies in C, i.e., if for any $x_1, x_2 \in C$ and any θ with $0 \leq \theta \leq 1$, we have
$$\theta x_1 + (1-\theta) x_2 \in C$$
4. Every affine set is also convex, since it contains the entire line between any two distinct points in it, and therefore also the line segment between the points.
5. A set C is called a **cone**, or nonnegative homogeneous, if for every $x \in C$ and $\theta \geq 0$ we have $\theta x \in C$.
6. A **hyperplane** is a set of the form
$$\left\{x \mid a^T x = b\right\}$$
where $a \in \mathbf{R}^n, a \neq 0$, and $b \in \mathbf{R}$.
7. Geometric interpretation of hyperplane:
$$\left\{x \mid a^T (x - x_0) = 0\right\}$$
where x_0 is any point in the hyperplane (*i.e.*, any point that satisfies $a^T x_0 = b$) This representation can in turn be expressed as
$$\left\{x \mid a^T (x - x_0) = 0\right\} = x_0 + a^{\perp}$$
where $a^{\perp}$ denotes the orthogonal complement of a. This shows that the hyperplane consists of an offset x_0, plus all vectors orthogonal to the (normal) vector a.
8. A **hyperplane** divides $\mathbf{R}^n$ into two halfspaces. A (closed) halfspace is a set of the form
$$\left\{x \mid a^T x \leq b\right\}$$
where $a \neq 0$. Geometric interpretation: the halfspace consists of x_0 plus any vector that makes an obtuse (or right) angle with the (outward normal) vector a.

9. A (Euclidean) **ball** (or just ball) in $\mathbf{R}^n$ has the form

$$B(x_c, r) = \{x \mid \|x - x_c\|_2 \le r\} = \left\{x \mid (x - x_c)^T (x - x_c) \le r^2\right\} = \{x_c + ru \mid \|u\|_2 \le 1\}$$

where $r > 0$, and $\|\cdot\|_2$ denotes the Euclidean norm, *i.e.*, $\|u\|_2 = \left(u^T u\right)^{1/2}$.

10. A Euclidean ball is a convex set.

11. A related family of convex sets is the **ellipsoids**, which have the form

$$\mathcal{E} = \left\{x \mid (x - x_c)^T P^{-1} (x - x_c) \le 1\right\}$$

where $P = P^T \succ 0$, i.e., P is symmetric and positive definite. The matrix P determines how far the ellipsoid extends in every direction from x_c; the lengths of the semi-axes of $\mathcal{E}$ are given by $\sqrt{\lambda_i}$, where λ_i are the eigenvalues of P. A ball is an ellipsoid with $P = r^2 I$. Since P is symmetric and positive definite, there exists u ($u^T = u^{-1}$) and Σ, s.t. $P = u^T \Sigma u$. Then

$$\begin{aligned}\mathcal{E} &= \left\{x \mid (x - x_c)^T P^{-1} (x - x_c) \le 1\right\} \\ &= \left\{x \mid (x - x_c)^T u^{-1} \Sigma^{-1} u^{-T} (x - x_c) \le 1\right\} \\ &= \left\{x \mid (u(x - x_c))^T \Sigma^{-1} (u(x - x_c)) \le 1\right\} \\ &= \left\{x \mid y^T \Sigma^{-1} y \le 1\right\} \\ &= \left\{x \mid \sum_{i=1}^{n} \frac{y_i^2}{r_i^2} \le 1\right\}\end{aligned}$$

12. Another common representation of an **ellipsoid** is

$$\mathcal{E} = \{x_c + Au \mid \|u\|_2 \le 1\}$$

where A is square and nonsingular. In this representation we can assume without loss of generality that A is symmetric and positive definite. By taking $A = P^{1/2}$ this representation gives the ellipsoid the form defined previously. When the matrix A is symmetric positive semidefinite but singular, the set is called a **degenerate ellipsoid**.

13. Ellipsoid's affine dimension is equal to the rank of A. Degenerate ellipsoids are also convex.

14. A norm ball of radius r and center x_c, given by $\{x \mid \|x - x_c\| \le r\}$, is convex.

15. The norm cone associated with the norm $\|\cdot\|$ is the set

$$C = \{(x, t) \mid \|x\| \le t\} \subseteq \mathbf{R}^{n+1}$$

16. The second-order cone (the quadratic cone, the Lorentz cone or ice-cream cone)

is the norm cone for the Euclidean norm, i.e.,

$$C = \left\{(x,t) \in \mathbf{R}^{n+1} \mid \|x\|_2 \le t\right\}$$
$$= \left\{ \begin{bmatrix} x \\ t \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix}^T \begin{bmatrix} I & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} \le 0, t \ge 0 \right\}$$

the second-order cone in $\mathbf{R}^3$ is $\left\{(x_1,x_2,t) \mid \left(x_1^2 + x_2^2\right)^{1/2} \le t\right\}$

17. A **polyhedron** is defined as the solution set of a finite number of linear equalities and inequalities:

$$\mathcal{P} = \left\{x \mid a_j^T x \le b_j, j = 1,\ldots,m, c_j^T x = d_j, j = 1,\ldots,p\right\}$$

A polyhedron is thus the intersection of a finite number of halfspaces and hyperplanes.

18. Affine sets (e.g., subspaces, hyperplanes, lines), rays, line segments, and halfspaces are all polyhedra.
19. polyhedra are convex sets.
20. A bounded polyhedron is sometimes called a **polytope**
21. Any notation for polyhedron

$$\mathcal{P} = \{x \mid Ax \preceq b, Cx = d\}$$

where

$$A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix}, \quad C = \begin{bmatrix} c_1^T \\ \vdots \\ c_p^T \end{bmatrix}$$

and the symbol $\preceq$ denotes vector inequality or componentwise inequality in $\mathbf{R}^m$:

$$u \preceq v \text{means} u_i \le v_i \text{for} i = 1,\ldots,m$$

22. The nonnegative **orthant** is the set of points with nonnegative components, *i.e.*

$$\mathbf{R}_+^n = \{x \in \mathbf{R}^n \mid x_i \ge 0, i = 1,\ldots,n\} = \{x \in \mathbf{R}^n \mid x \succeq 0\}$$

23. The nonnegative orthant is a polyhedron and a cone (and therefore called a polyhedral cone)

24. Suppose the $k+1$ points $v_0, \ldots, v_k \in \mathbf{R}^n$ are affinely independent, which means $v_1 - v_0, \ldots, v_k - v_0$ are linearly independent. The simplex determined by them is given by

$$C = \operatorname{conv}\{v_0, \ldots, v_k\} = \left\{\theta_0 v_0 + \cdots + \theta_k v_k \mid \theta \succeq 0, \mathbf{1}^T \theta = 1\right\}$$

25. The affine dimension of this simplex is k, so it is sometimes referred to as a k -dimensional simplex in $\mathbf{R}^n$.
26. A 1-dimensional simplex is a line segment; a 2-dimensional simplex is a triangle (including its interior); and a 3-dimensional simplex is a tetrahedron. The unit simplex is the n -dimensional simplex determined by the zero vector and the unit vectors, *i.e.*,$0, e_1, \ldots, e_n \in \mathbf{R}^n$. It can be expressed as the set of vectors that satisfy

$$x \succeq 0, \quad \mathbf{1}^T x \leq 1$$

27. To describe the simplex as a polyhedron, see P33 of Boyd.
28. $\mathbf{S}^n$ to denote the set of symmetric $n \times n$ matrices,

$$\mathbf{S}^n = \left\{X \in \mathbf{R}^{n \times n} \mid X = X^T\right\}$$

which is a vector space with dimension $n(n+1)/2$.
29. $\mathbf{S}_+^n$ to denote the set of symmetric positive semidefinite matrices:

$$\mathbf{S}_+^n = \{X \in \mathbf{S}^n \mid X \succeq 0\}$$

30. $\mathbf{S}_{++}^n+$ to denote the set of symmetric positive definite matrices:

$$\mathbf{S}_{++}^n = \{X \in \mathbf{S}^n \mid X \succ 0\}$$

31. The set $\mathbf{S}_+^n$ is a convex cone.
32. A function $f : \mathbf{R}^n \to \mathbf{R}^m$ is affine if it is a sum of a linear function and a constant, i.e., if it has the form $f(x) = Ax + b$, where $A \in \mathbf{R}^{m \times n}$ and $b \in \mathbf{R}^m$.
33. Suppose $S \subseteq \mathbf{R}^n$ is convex and $f : \mathbf{R}^n \to \mathbf{R}^m$ is an affine function. Then the image of S under f

$$f(S) = \{f(x) \mid x \in S\}$$

is convex.

34. Similarly, if $f : \mathbf{R}^k \to \mathbf{R}^n$ is an affine function, the inverse image of S under f

$$f^{-1}(S) = \{x \mid f(x) \in S\}$$

is convex.

35. Convex set's scaling and translation are still convex sets.

36. The projection of a convex set onto some of its coordinates is convex: if $S \subseteq \mathbf{R}^m \times \mathbf{R}^n$ is convex, then

$$T = \{x_1 \in \mathbf{R}^m \mid (x_1, x_2) \in S \text{ for some } x_2 \in \mathbf{R}^n\}$$

is convex.

37. The sum of two sets is defined as

$$S_1 + S_2 = \{x + y \mid x \in S_1, y \in S_2\}$$

If S_1 and S_2 are convex, then $S_1 + S_2$ is convex. To see this, if S_1 and S_2 are convex, then so is the direct or Cartesian product

$$S_1 \times S_2 = \{(x_1, x_2) \mid x_1 \in S_1, x_2 \in S_2\}$$

The image of this set under the linear function $f(x_1, x_2) = x_1 + x_2$ is the sum

$$S_1 + S_2$$

38. We can also consider the partial sum of $S_1, S_2 \in \mathbf{R}^n \times \mathbf{R}^m$, defined as

$$S = \{(x, y_1 + y_2) \mid (x, y_1) \in S_1, (x, y_2) \in S_2\}$$

where $x \in \mathbf{R}^n$ and $y_i \in \mathbf{R}^m$. For $m = 0$, the partial sum gives the intersection of S_1 and S_2; for $n = 0$, it is set addition. Partial sums of convex sets are convex.

39. The perspective function $P : \mathbf{R}^{n+1} \to \mathbf{R}^n$, with domain $\operatorname{dom} P = \mathbf{R}^n \times \mathbf{R}_{++}$, as $P(z, t) = z/t$.

40. If $C \subseteq \operatorname{dom} P$ is convex, then its image

$$P(C) = \{P(x) \mid x \in C\}$$

is convex.

41. The inverse image of a convex set under the perspective function is also convex: if $C \subseteq \mathbf{R}^n$ is convex, then

$$P^{-1}(C) = \left\{(x, t) \in \mathbf{R}^{n+1} \mid x/t \in C, t > 0\right\}$$

is convex.

42. A cone $K \subseteq \mathbf{R}^n$ is called a **proper cone** if it satisfies the following:
    (a) K is convex.
    (b) K is closed.
    (c) K is solid, which means it has nonempty interior.
    (d) K is pointed, which means that it contains no line (or equivalently, $x \in$

$$K, -x \in K \Longrightarrow x = 0)$$

43. Associate with the proper cone K the **partial ordering** on $\mathbf{R}^n$ defined by

$$x \preceq_K y \Longleftrightarrow y - x \in K$$

44. Similarly, we define an associated **strict partial ordering** by

$$x \prec_K y \Longleftrightarrow y - x \in \operatorname{int} K$$

    and write $x \succ_K y$ for $y \prec_K x$.
45. When $K = \mathbf{R}_+$, the partial ordering $\preceq_K$ is the usual ordering $\leq$ on $\mathbf{R}$, and the strict partial ordering $\prec_K$ is the same as the usual strict ordering $<$ on $\mathbf{R}$.
46. Component-wise inequality between vectors: $x \preceq_K y$ means that $x_i \leq y_i$ $i = 1, \ldots, n$. Usual matrix inequality: $X \preceq_K Y$ means $Y - X$ is positive semidefinite.
47. A generalized inequality $\preceq_K$ satisfies many properties, such as
    (a) $\preceq_K$ is preserved under addition: if $x \preceq_K y$ and $u \preceq_K v$, then $x + u \preceq_K y + v$
    (b) $\preceq_K$ is transitive: if $x \preceq_K y$ and $y \preceq_K z$ then $x \preceq_K z$
    (c) $\preceq_K$ is preserved under nonnegative scaling: if $x \preceq_K y$ and $\alpha \geq 0$ then $\alpha x \preceq K \alpha y$
    (d) $\preceq_K$ is reflexive: $x \preceq Kx$
    (e) $\preceq_K$ is antisymmetric: if $x \preceq_K y$ and $y \preceq_K x$, then $x = y$
    (f) $\preceq_K$ is preserved under limits: if $x_i \preceq_K y_i$ for $i = 1, 2, \ldots, x_i \to x$ and $y_i \to y$ as $i \to \infty$, then $x \preceq Ky$

    The corresponding strict generalized inequality $\prec_K$ satisfies, for example,
    (a) if $x \prec_K y$ then $x \preceq_K y$
    (b) if $x \prec_K y$ and $u \preceq_K v$ then $x + u \prec_K y + v$
    (c) if $x \prec_K y$ and $\alpha > 0$ then $\alpha x \prec_K \alpha y$
    (d) $x \not\prec_K x$
    (e) if $x \prec_K y$, then for u and v small enough, $x + u \prec_K y + v$
48. $x \in S$ is the **minimum** element of S if for every $y \in S$ we have $x \preceq_K y$. (means any point could be compared with minimum)
49. If a set has a **minimum** (maximum) element, then it is unique.

50. We say that $x \in S$ is a **minimal** element of S (with respect to the generalized inequality $\preceq_K$) if $y \in S, y \preceq_K x$ only if $y = x$.
51. A set can have many different minimal (maximal) elements.
52. A point $x \in S$ is the **minimum** element of S if and only if

$$S \subseteq x + K$$

Here $x + K$ denotes all the points that are comparable to x and greater than or equal to x (according to $\preceq_K$).
53. A point $x \in S$ is a **minimal** element if and only if

$$(x - K) \cap S = \{x\}$$

Here $x - K$ denotes all the points that are comparable to x and less than or equal to x (according to $\preceq_K$); the only point in common with S is x.

Definition 1.2.2 — Linear Program. The optimization problem is called a linear program if the objective and constraint functions $f_0, \dots, f_m$ are linear, i.e., satisfy

$$f_i(\alpha x + \beta y) = \alpha f_i(x) + \beta f_i(y)$$

for all $x, y \in \mathbf{R}^n$ and all $\alpha, \beta \in \mathbf{R}$.

Definition 1.2.3 — Convex Optimization Problem. A **convex optimization problem** is one in which the objective and constraint functions are convex, which means they satisfy the inequality

$$f_i(\alpha x + \beta y) \leq \alpha f_i(x) + \beta f_i(y)$$

for all $x, y \in \mathbf{R}^n$ and all $\alpha, \beta \in \mathbf{R}$ with $\alpha + \beta = 1, \alpha \geq 0, \beta \geq 0$.

we see that convexity is more general than linearity.

1.2.1 Matrix derivation

Definition 1.2.4
1. one variable: $df = f'(x)dx$
2. multivariables: $df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i = \frac{\partial f}{\partial x}^T d\boldsymbol{x}$,
3. matrix: $df = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\partial f}{\partial X_{ij}} dX_{ij} = \operatorname{tr}\left(\frac{\partial f}{\partial X}^T dX\right)$

Proposition 1.2.1 — Properties of Matrix derivation.
1. $d(X \pm Y) = dX \pm dY$
2. $d(XY) = (dX)Y + XdY$
3. $d\left(X^T\right) = (dX)^T$;
4. $d\operatorname{tr}(X) = \operatorname{tr}(dX)$
5. $dX^{-1} = -X^{-1}dXX^{-1}$ (derivation at both sides of $XX^{-1} = I$)

6. $d|X| = \operatorname{tr}(X^* dX)$
7. $d|X| = |X| \operatorname{tr}\left(X^{-1} dX\right)$ (when X is inversible)
8. $d(X \odot Y) = dX \odot Y + X \odot dY$ ($\odot$ is a pointwise operator)
9. $d\sigma(X) = \sigma'(X) \odot dX, \quad \sigma(X) = \left[\sigma\left(X_{ij}\right)\right]$ is a pointwise function operator, $\sigma'(X) = \left[\sigma'\left(X_{ij}\right)\right]$ is pointwise function, such as

$$X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}, d\sin(X) = \begin{bmatrix} \cos X_{11} dX_{11} & \cos X_{12} dX_{12} \\ \cos X_{21} dX_{21} & \cos X_{22} dX_{22} \end{bmatrix} = \cos(X) \odot dX$$

10. Chain rule: Suppose $Y = AXB$, then $df = \operatorname{tr}\left(\frac{\partial f^T}{\partial Y} dY\right) = \operatorname{tr}\left(\frac{\partial f^T}{\partial Y} \operatorname{Ad} XB\right) = \operatorname{tr}\left(B \frac{\partial f^T}{\partial Y} AdX\right) = \operatorname{tr}\left(\left(A^T \frac{\partial f}{\partial Y} B^T\right)^T dX\right)$, so $\frac{\partial f}{\partial X} = A^T \frac{\partial f}{\partial Y} B^T$. Where $dY = (dA)XB + AdXB + AXdB = AdXB$. Since A, B are constant, $dA = 0, dB = 0$, so $\frac{\partial f^T}{\partial Y} AdX$ and B

Proposition 1.2.2 — Some properties of trace. 1. constant: $a = \operatorname{tr}(a)$
2. $\operatorname{tr}\left(A^T\right) = \operatorname{tr}(A)$
3. $\operatorname{tr}(A \pm B) = \operatorname{tr}(A) \pm \operatorname{tr}(B)$
4. $\operatorname{tr}(AB) = \operatorname{tr}(BA) = \sum_{i,j} A_{ij} B_{ji}$ where the sizes of A and B^T should be same
5. $\operatorname{tr}\left(A^T(B \odot C)\right) = \operatorname{tr}\left((A \odot B)^T C\right) = \sum_{i,j} A_{ij} B_{ij} C_{ij}$ where the sizes of A, B, C should be the same

■ **Example 1.1** $l = \|X\boldsymbol{w} - \boldsymbol{y}\|^2$, What is the LSE of $\boldsymbol{w}$, which is the zero point of $\frac{\partial l}{\partial \boldsymbol{w}}$. Where $\boldsymbol{y}$ is $m \times 1$ column vector, X is $m \times n$ matrix, $\boldsymbol{w}$ is $n \times 1$ column vector, l is a number.

Proof. $l = (X\boldsymbol{w} - \boldsymbol{y})^T (X\boldsymbol{w} - \boldsymbol{y})$, $dl = (Xd\boldsymbol{w})^T (X\boldsymbol{w} - \boldsymbol{y}) + (X\boldsymbol{w} - \boldsymbol{y})^T (Xd\boldsymbol{w}) = 2(X\boldsymbol{w} - \boldsymbol{y})^T Xd\boldsymbol{w}$, where $Xd\boldsymbol{w}$ and $X\boldsymbol{w} - \boldsymbol{y}$ are two vectors, and they satisfies $\boldsymbol{u}^T\boldsymbol{v} = \boldsymbol{v}^T\boldsymbol{u}$. According to the relationship between derivatives and differentiation $dl = \frac{\partial l}{\partial \boldsymbol{w}}^T d\boldsymbol{w}$, we obtain $\frac{\partial l}{\partial \boldsymbol{w}} = 2X^T(X\boldsymbol{w} - \boldsymbol{y})$. Let $\frac{\partial l}{\partial \boldsymbol{w}} = \boldsymbol{0}$, which $X^T X\boldsymbol{w} = X^T\boldsymbol{y}$, and the LSE of $\boldsymbol{w}$ is $\boldsymbol{w} = \left(X^T X\right)^{-1} X^T \boldsymbol{y}$ ■

1.2.2 Least-squares problems

Definition 1.2.5 A least-squares problem is an optimization problem with no constraints (i . e ., m=0) and an objective which is a sum of squares of terms of the form $a_i^T x - b_i$

$$\text{minimize} \quad f_0(x) = \|Ax - b\|_2^2 = \sum_{i=1}^{k} \left(a_i^T x - b_i\right)^2$$

Here $A \in \mathbf{R}^{k \times n}$(with $k \geq n$), a_i^T are the rows of A, and the vector $x \in \mathbf{R}^n$ is the optimization variable.

The solution of a least-squares problem can be reduced to solving a set of linear equations, just multify A^T at both sides:

$$\left(A^T A\right) x = A^T b$$

so we have the analytical solution $x = \left(A^T A\right)^{-1} A^T b$.

Proposition 1.2.3 — Least-squares problems's proposition.

1. The least-squares problem can be solved in a time approximately proportional to $n^2 k$, with a known constant.
2. The matrix A is sparse, which means that it has far fewer than kn nonzero entries.
3. Recognizing an optimization problem: the objective is a quadratic function and then test whether the associated quadratic form is positive semidenite.
4. weighted least-squares: $\sum w_i (a_i^T x - b_i)^2$
5. Regularization: extra terms are added to the cost function

Definition 1.2.6 — Linear programming. Linear programming: the objective and all constraint functions are linear:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, m \end{array}$$

Many questions can be transformed to linear program, such as the Chebyshev approximation problem.

Definition 1.2.7 — Convex optimization. A convex optimization problem is one of the form

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq b_i, \quad i = 1, \ldots, m \end{array}$$

where the functions $f_0, \ldots, f_m : \mathbf{R}^n \to \mathbf{R}$ are convex, *i.e.*, satisfy

$$f_i(\alpha x + \beta y) \leq \alpha f_i(x) + \beta f_i(y)$$

for all $x, y \in \mathbf{R}^n$ and all $\alpha, \beta \in \mathbf{R}$ with $\alpha + \beta = 1, \alpha \geq 0, \beta \geq 0$.

Solving convex optimization problems: interior-point methods.

R The least-squares problem and linear programming problem are both special cases of the general convex optimization problem.

Definition 1.2.8 — Nonlinear optimization.

1. For nonlinear optimization, the locally optimal point is more important: it minimizes the objective function among feasible points that are near it, but is not guaranteed to have a lower objective value than all other feasible points.
2. Locally optimal point's criterion: dierentiability of the objective and constraint

functions.

3. Disadvantage of local optimal method:
   (a) For locally optimal point: This initial guess or starting point is critical, and can greatly aect the objective value of the local solution obtained.
   (b) Local optimization methods are often sensitive to algorithm parameter values, which may need to be adjusted for a particular problem, or family of problems.
4. Role of convex optimization in nonconvex problems
   (a) Initialization for local optimization: Starting with a nonconvex problem, we rst nd an approximate, but convex, formulation of the problem. By solving this approximate problem, which can be done easily and without an initial guess, we obtain the exact solution to the approximate convex problem. This point is then used as the starting point.
   (b) Bounds for global optimization: In relaxation, each nonconvex constraint is replaced with a looser, but convex, constraint.

1.2.3 Convex Set

Theorem 1.2.4 If C is an affine set and $x_0 \in C$, then the set

$$V = C - x_0 = \{x - x_0 \mid x \in C\}$$

is a subspace, *i.e.*, closed under sums and scalar multiplication. Thus, the affine set C can be expressed as

$$C = V + x_0 = \{v + x_0 \mid v \in V\}$$

i.e., as a subspace plus an offset. The subspace V associated with the affine set C does not depend on the choice of x_0, so x_0 can be chosen as any point in C. We define the dimension of an affine set C as the dimension of the subspace $V = C - x_0$ where x_0 is any element of C.

Definition 1.2.9 — Affine hull. The set of all affine combinations of points in some set $C \subseteq \mathbf{R}^n$ is called the affine hull of C, and denoted aff C :

$$\mathbf{aff} C = \{\theta_1 x_1 + \cdots + \theta_k x_k \mid x_1, \ldots, x_k \in C, \theta_1 + \cdots + \theta_k = 1\}$$

The affine hull is the smallest affine set that contains C.

Definition 1.2.10 If the affine dimension of a set $C \subseteq \mathbf{R}^n$ is less than n, then the set lies in the affine set aff $C \neq \mathbf{R}^n$.

1. We define the relative interior of the set C, denoted relint C, as its interior relative

to aff C :

$$\mathrm{relint} C = \{x \in C \mid B(x,r) \cap \mathrm{aff}\, C \subseteq C \text{ for some } r > 0\}$$

where $B(x,r) = \{y \mid \|y - x\| \leq r\}$, the ball of radius r and center x in the norm $\|\cdot\|$ · (Here $\|\cdot\|$ is any norm; all norms define the same relative interior.)

2. We can then define the relative boundary of a set C as $\mathrm{cl}\, C \setminus$ relint C, where $\mathrm{cl}\, C$ is the closure of C.

Definition 1.2.11 — Convex Hull. The **convex hull** of a set C, denoted conv C, is the set of all convex combinations of points in C :

$$\mathrm{conv}\, C = \{\theta_1 x_1 + \cdots + \theta_k x_k \mid x_i \in C, \theta_i \geq 0, i = 1, \ldots, k, \theta_1 + \cdots + \theta_k = 1\}$$

1. The convex hull conv C is always convex.
2. It is the smallest convex set that contains C
3. Any subspace is ane, and a convex cone (hence convex).

Definition 1.2.12 — Conic Hull. The **conic hull** of a set C is the set of all conic combinations of points in C, i.e.,

$$\{\theta_1 x_1 + \cdots + \theta_k x_k \mid x_i \in C, \theta_i \geq 0, i = 1, \ldots, k\}$$

which is also the smallest convex cone that contains C .

Theorem 1.2.5 — Some operations that preserve the convexty.

1. Intersection: if S_α is convex for every $\alpha \in \mathcal{A}$, then $\bigcap_{\alpha \in \mathcal{A}} S_\alpha$ is convex. (Subspaces, affine sets, and convex cones are also closed under arbitrary intersections.)
2. Conversely, every closed convex set S is a (usually infinite) intersection of halfspaces. The closed convex set S is the intersection of all halfspaces that contain it:

$$S = \bigcap \{\mathcal{H} \mid \mathcal{H} \text{ halfspace}, S \subseteq \mathcal{H}\}$$

3. A **linear-fractional function** is formed by composing the perspective function with an affine function. Suppose $g : \mathbf{R}^n \to \mathbf{R}^{m+1}$ is affine, *i.e.*

$$g(x) = \begin{bmatrix} A \\ c^T \end{bmatrix} x + \begin{bmatrix} b \\ d \end{bmatrix}$$

where $A \in \mathbf{R}^{m \times n}, b \in \mathbf{R}^m, c \in \mathbf{R}^n$, and $d \in \mathbf{R}$. The function $f : \mathbf{R}^n \to \mathbf{R}^m$ given by

$f = P \circ g$, i.e.

$$f(x) = (Ax+b)/\left(c^T x + d\right), \quad \mathrm{dom}\, f = \left\{x \mid c^T x + d > 0\right\}$$

is called a linear-fractional (or projective) function. If $c = 0$ and $d > 0$, the domain of f is $\mathbf{R}^n$, and f is an affine function. So we can think of affine and linear functions as special cases of linear-fractional functions.

4. **Projective interpretation**: represent a linearfractional function as a matrix

$$Q = \begin{bmatrix} A & b \\ c^T & d \end{bmatrix} \in \mathbf{R}^{(m+1)\times(n+1)}$$

that acts on (multiplies) points of form $(x,1)$, which yields $\left(Ax+b, c^T x + d\right)$. This result is then scaled or normalized so that its last component is one, which yields

$$(f(x), 1)$$

The linear-fractional function (2.13) can be expressed as

$$f(x) = \mathcal{P}^{-1}(Q\mathcal{P}(x))$$

Thus, we start with $x \in \mathrm{dom}\, f$, i.e., $c^T x + d > 0$. We then form the ray $\mathcal{P}(x)$ in $\mathbf{R}^{n+1}$. The linear transformation with matrix Q acts on this ray to produce another ray $Q\mathcal{P}(x)$. since $x \in \mathrm{dom}\, f$, the last component of this ray assumes positive values. Finally we take the inverse projective transformation to recover $f(x)$. linear-fractional functions preserve convexity.

Theorem 1.2.6 — separating hyperplane theorem. Suppose C and D are nonempty disjoint convex sets, i.e.,$C \cap D = \emptyset$. Then there exist $a \neq 0$ and b such that $a^T x \leq b$ for all $x \in C$ and $a^T x \geq b$ for all $x \in D$. In other words, the affine function $a^T x - b$ is nonpositive on C and nonnegative on D. The hyperplane $\left\{x \mid a^T x = b\right\}$ is called a **separating hyperplane** for the sets C and D, or is said to separate the sets C and D.

R [strict separation] The stronger condition that $a^T x < b$ for all $x \in C$ and $a^T x > b$ for all $x \in D$. This is called strict separation of the sets C and D. (for example, one set is open set and the other is closed.)

Proof. Here we consider a special case, and leave the extension of the proof to the general case as an exercise (exercise 2.22). We assume that the (Euclidean) distance between C and D, defined as

$$\mathrm{dist}(C,D) = \inf\{\|u - v\|_2 \mid u \in C, v \in D\}$$

is positive (two sets are disjoint), and that there exist points $c \in C$ and $d \in D$ that achieve the minimum distance, i.e., $\|c-d\|_2 = \mathrm{dist}(C,D)$. (These conditions are satisfied, for example, when C and D are closed and one set is bounded.) Define

$$a = d - c, \quad b = \frac{\|d\|_2^2 - \|c\|_2^2}{2}$$

We will show that the affine function

$$f(x) = a^T x - b = (d-c)^T (x - (1/2)(d+c))$$

is nonpositive on C and nonnegative on D, i.e., that the hyperplane $\{x \mid a^T x = b\}$ separates C and D. This hyperplane is perpendicular to the line segment between c and d, and passes through its midpoint.

We first show that f is nonnegative on D. The proof that f is nonpositive on C is similar (or follows by swapping C and D and considering $-f$). Suppose there were a point $u \in D$ for which

$$f(u) = (d-c)^T (u - (1/2)(d+c)) < 0 \tag{1.3}$$

We can express $f(u)$ as

$$f(u) = (d-c)^T (u - d + (1/2)(d-c)) = (d-c)^T (u-d) + (1/2)\|d-c\|_2^2 < 0$$

We see that $(d-c)^T(u-d) < 0$. Now we observe that

$$\left.\frac{d}{dt}\|d + t(u-d) - c\|_2^2\right|_{t=0} = 2(d-c)^T(u-d) < 0$$

so for some small $t > 0$, with $t \leq 1$, we have

$$\|d + t(u-d) - c\|_2 < \|d-c\|_2$$

i.e., the point $d + t(u-d)$ is closer to c than d is. Since D is convex and contains d and u, we have $d + t(u-d) \in D$. But this is impossible, since d is assumed to be the point in D that is closest to C. ■

■ Example 1.2 — Separation of an affine and a convex set. Suppose C is convex and D is affine, *i.e.*, $D = \{Fu + g \mid u \in \mathbf{R}^m\}$, where $F \in \mathbf{R}^{n\times m}$. Suppose C and D are disjoint, so by the separating hyperplane theorem there are $a \neq 0$ and b such that $a^T x \leq b$ for all $x \in C$ and $a^T x \geq b$ for all $x \in D$. Now $a^T x \geq b$ for all $x \in D$ means $a^T F u \geq b - a^T g$ for all $u \in \mathbf{R}^m$. But **a linear function is bounded below on $\mathbf{R}^m$ only when it is zero**, so we conclude $a^T F = 0$ (and hence, $b \leq a^T g$). Thus we conclude that there exists $a \neq 0$ such that $F^T a = 0$ and $a^T x \leq a^T g$ for all $x \in C$

■ **Example 1.3 — Strict separation of a point and a closed convex set.** Let C be a closed convex set and $x_0 \notin C$. Then there exists a hyperplane that strictly separates x_0 from C

To see this, note that the two sets C and $B(x_0, \epsilon)$ do not intersect for some $\epsilon > 0$. By the separating hyperplane theorem, there exist $a \neq 0$ and b such that $a^T x \leq b$ for $x \in C$ and $a^T x \geq b$ for $x \in B(x_0, \epsilon)$.

Using $B(x_0, \epsilon) = \{x_0 + u \mid \|u\|_2 \leq \epsilon\}$, the second condition can be expressed as

$$a^T(x_0 + u) \geq b \text{ for all } \|u\|_2 \leq \epsilon$$

The u that minimizes the lefthand side is $u = -\epsilon a / \|a\|_2$;(How to solve the minimum of $a^T u$ under $\|u\|_2 < \varepsilon$) using this value we have

$$a^T x_0 - \epsilon\|a\|_2 \geq b$$

Therefore the affine function

$$f(x) = a^T x - b - \epsilon\|a\|_2/2$$

is negative on C and positive at x_0.

Corollary 1.2.7 From previous example, we can obtain a conclusion: a closed convex set is the intersection of all halfspaces that contain it.

Proof. Indeed, let C be closed and convex, and let S be the intersection of all halfspaces containing C. Obviously $x \in C \Rightarrow x \in S$. To show the converse, suppose there exists $x \in S, x \notin C$. By the strict separation result there exists a hyperplane that strictly separates x from C, i.e., there is a halfspace containing C but not x. In other words, $x \notin S$. ■

Theorem 1.2.8 — Theorem of alternatives for strict linear inequalities. We derive the necessary and sufficient conditions for solvability of a system of strict linear inequalities

$$Ax \prec b \tag{1.4}$$

These inequalities are infeasible if and only if the (convex) sets

$$C = \{b - Ax \mid x \in \mathbf{R}^n\}, \quad D = \mathbf{R}^m_{++} = \{y \in \mathbf{R}^m \mid y \succ 0\} \tag{1.5}$$

do not intersect. The set D is open; C is an affine set. Hence by the result above, C and D are disjoint if and only if there exists a separating hyperplane, i.e., a nonzero $\lambda \in \mathbf{R}^m$ and $\mu \in \mathbf{R}$ such that $\lambda^T y \leq \mu$ on C and $\lambda^T y \geq \mu$ on D.

Each of these conditions can be simplified. The first means $\lambda^T(b - Ax) \leq \mu$ for all x. This implies (as in example 1.2) that $A^T\lambda = 0$ and $\lambda^T b \leq \mu$. The second inequality means $\lambda^T y \geq \mu$ for all $y \succ 0$. This implies $\mu \leq 0$ and $\lambda \succeq 0, \lambda \neq 0$

Putting it all together, we find that the set of strict inequalities (1.5) is infeasible if and only if there exists $\lambda \in \mathbf{R}^m$ such that

$$\lambda \neq 0, \quad \lambda \succeq 0, \quad A^T\lambda = 0, \quad \lambda^T b \leq 0 \tag{1.6}$$

This is also a system of linear inequalities and linear equations in the variable $\lambda \in \mathbf{R}^m$. We say that 1.5 and (1.6) form a pair of alternatives: for any data A and b, exactly one of them is solvable.

1.2.4 Support hyperplane

Definition 1.2.13 — Supporting hyperplanes. Suppose $C \subseteq \mathbf{R}^n$, and x_0 is a point in its boundary bd C, i.e.,

$$x_0 \in \mathbf{bd}C = \mathbf{cl}C \setminus \text{int } C$$

If $a \neq 0$ satisfies $a^T x \leq a^T x_0$ for all $x \in C$, then the hyperplane $\{x \mid a^T x = a^T x_0\}$ is called a **supporting hyperplane** to C at the point x_0. This is equivalent to saying that the hyperplane $\{x \mid a^T x = a^T x_0\}$ is tangent to C at x_0, and the halfspace $\{x \mid a^T x \leq a^T x_0\}$ contains C.

Theorem 1.2.9 — Supporting hyperplane theorem. For any nonempty convex set C, and any $x_0 \in$ bd C, there exists a supporting hyperplane to C at x_0.

Proof. We distinguish two cases. If the interior of C is nonempty, the result follows immediately by applying the separating hyperplane theorem to the sets $\{x_0\}$ and int C. If the interior of C is empty, then C must lie in an affine set of dimension less than n, and any hyperplane containing that affine set contains C and x_0, and is a (trivial) supporting hyperplane. ∎

There is also a partial converse of the supporting hyperplane theorem: If a set is closed, has nonempty interior, and has a supporting hyperplane at every point in its boundary, then it is convex. (See exercise 2.27.)

1.2.5 Dual Cone

Definition 1.2.14 — Dual cones. Let K be a cone. The set

$$K^* = \left\{y \mid x^T y \geq 0 \text{ for all } x \in K\right\}$$

is called the **dual cone** of K. As the name suggests, K^* is a cone, and is always convex, even when the original cone K is not.

Geometrically, $y \in K^*$ if and only if $-y$ is the normal of a hyperplane that supports K at the origin.

■ **Example 1.4 — Subspace.** The dual cone of a subspace $V \subseteq \mathbf{R}^n$ (which is a cone) is its orthogonal complement $V^\perp = \{y \mid v^T y = 0 \text{ for all } v \in V\}$

■ **Example 1.5 — Nonnegative orthant.** The cone $\mathbf{R}^n_+$ is its own dual:

$$x^T y \geq 0 \text{ for all } x \succeq 0 \Longleftrightarrow y \succeq 0$$

We call such a **cone self-dual**.

■ **Example 1.6 — Positive semidefinite cone.** On the set of symmetric $n \times n$ matrices $\mathbf{S}^n$, we use the standard inner product $\operatorname{tr}(XY) = \sum_{i,j=1}^n X_{ij} Y_{ij}$. The positive semidefinite cone $\mathbf{S}^n_+$ is self-dual, i.e., for $X, Y \in \mathbf{S}^n$,

$$\operatorname{tr}(XY) \geq 0 \text{ for all } X \succeq 0 \Longleftrightarrow Y \succeq 0$$

We will establish this fact. Suppose $Y \notin \mathbf{S}^n_+$. Then there exists $q \in \mathbf{R}^n$ with

$$q^T Y q = \operatorname{tr}\left(q q^T Y\right) < 0$$

Hence the positive semidefinite matrix $X = qq^T$ satisfies $\operatorname{tr}(XY) < 0$; it follows that $Y \notin (\mathbf{S}^n_+)^*$

Now suppose $X, Y \in \mathbf{S}^n_+$. We can express X in terms of its **eigenvalue decomposition** as $X = \sum_{i=1}^n \lambda_i q_i q_i^T$, where (the eigenvalues) $\lambda_i \geq 0, i = 1, \ldots, n$. Then we have

$$\operatorname{tr}(YX) = \operatorname{tr}\left(Y \sum_{i=1}^n \lambda_i q_i q_i^T\right) = \sum_{i=1}^n \lambda_i q_i^T Y q_i \geq 0$$

This shows that $Y \in (\mathbf{S}^n_+)^*$

■ **Example 1.7 — Dual of a norm cone.** Let $\|\cdot\|$ be a norm on $\mathbf{R}^n$. The dual of the associated cone $K = \{(x,t) \in \mathbf{R}^{n+1} \mid \|x\| \leq t\}$ is the cone defined by the dual norm, i.e.

$$K^* = \left\{(u,v) \in \mathbf{R}^{n+1} \mid \|u\|_* \leq v\right\}$$

where the dual norm is given by $\|u\|_* = \sup\{u^T x \mid \|x\| \leq 1\}$

To prove the result we have to show that

$$(x,t)^T \cdot (u,v) = x^T u + tv \geq 0 \text{ whenever } \|x\| \leq t \Longleftrightarrow \|u\|_* \leq v$$

Let us start by showing that the righthand condition on (u,v) implies the lefthand condition. Suppose $\|u\|_* \leq v$, and $\|x\| \leq t$ for some $t > 0$. (If $t = 0, x$ must be zero, so obviously $u^T x + vt \geq 0$.) Applying the definition of the dual norm, and the fact that $\|-x/t\| \leq 1$, we have

$$u^T(-x/t) \leq \|u\|_* \leq v$$

and therefore $u^T x + vt \geq 0$.

Next we show that the lefthand condition implies the righthand condition. Suppose $\|u\|_* > v$, i.e., that the righthand condition does not hold. Then by the definition of the dual norm, there exists an x with $\|x\| \leq 1$ and $x^T u > v$. Taking $t = 1$, we have

$$u^T(-x) + v < 0$$

which contradicts the lefthand condition.

Proposition 1.2.10 — Properties of Dual cones. 1. K^* is closed and convex.

2. $K_1 \subseteq K_2$ implies $K_2^* \subseteq K_1^*$
3. If K has nonempty interior, then K^* is pointed.
4. If the closure of K is pointed then K^* has nonempty interior.
5. K^{**} is the closure of the convex hull of K. (Hence if K is convex and closed, $K^{**} = K$.)

These properties show that if K is a proper cone, then so is its dual K^*, and moreover, that $K^{**} = K$

Definition 1.2.15 — Generalized inequality. Now suppose that the convex cone K is proper, so it induces a generalized inequality $\preceq_K$. Then its dual cone K^* is also proper, and therefore induces a generalized inequality. We refer to the generalized inequality $\preceq_{K^*}$ as the *dual* of the generalized inequality $\preceq_K$

Proposition 1.2.11 — Properties of generalized inequality. Some important properties relating a generalized inequality and its dual are:

1. $\preceq_K y$ if and only if $\lambda^T x \leq \lambda^T y$ for all $\lambda \succeq_K \cdot 0$
2. $x \prec_K y$ if and only if $\lambda^T x < \lambda^T y$ for all $\lambda \succeq_K \cdot 0, \lambda \neq 0$

R Since $K = K^{**}$, the dual generalized inequality associated with $\preceq_{K^*}$ is $\preceq_K$, so these properties hold if the generalized inequality and its dual are swapped. As a specific example, we have $\lambda \preceq_{K^*} \mu$ if and only if $\lambda^T x \leq \mu^T x$ for all $x \succeq \kappa 0$

Theorem 1.2.12 — Theorem of alternatives for linear strict generalized inequalities.. Suppose $K \subseteq \mathbf{R}^m$ is a proper cone. Consider the strict generalized inequality

$$Ax \prec_K b \tag{1.7}$$

where $x \in \mathbf{R}^n$. Suppose it is infeasible, i.e., the affine set $\{b - Ax \mid x \in \mathbf{R}^n\}$ does not intersect the open convex set int K. Then there is a separating hyperplane, *i.e.*, a nonzero $\lambda \in \mathbf{R}^m$ and $\mu \in \mathbf{R}$ such that $\lambda^T(b - Ax) \leq \mu$ for all x, and $\lambda^T y \geq \mu$ for all $y \in \operatorname{int} K$. The first condition implies $A^T\lambda = 0$ and $\lambda^T b \leq \mu$. The second condition implies $\lambda^T y \geq \mu$ for all $y \in K$, which can only happen if $\lambda \in K^*$ and $\mu \leq 0$.

Putting it all together we find that if (2.21) is infeasible, then there exists λ such that

$$\lambda \neq 0, \quad \lambda \succeq_{K^*} 0, \quad A^T\lambda = 0, \quad \lambda^T b \leq 0 \tag{1.8}$$

Now we show the converse: if 1.8 holds, then the inequality system 1.7 cannot be feasible. Suppose that both inequality systems hold. Then we have $\lambda^T(b - Ax) > 0$, since $\lambda \neq 0, \lambda \succeq_{K^*} 0$, and $b - Ax \succ_K 0$. But using $A^T\lambda = 0$ we find that $\lambda^T(b - Ax) = \lambda^T b \leq 0$, which is a contradiction.

Thus, the inequality systems 1.7 and 1.8 are alternatives: for any data A, b, exactly one of them is feasible. (This generalizes the alternatives 1.5 ,1.6 for the special case $K = \mathbf{R}^m_+$.)

Definition 1.2.16 — Minimum element. x is the minimum element of S, with respect to the generalized inequality $\preceq_K$, if and only if for all $\lambda \succ_{K^*} 0, x$ is the unique minimizer of $\lambda^T z$ over $z \in S$. Geometrically, this means that for any $\lambda \succ_{K^*} 0$, the hyperplane

$$\left\{z \mid \lambda^T(z - x) = 0\right\}$$

is a strict supporting hyperplane to S at x. (By strict supporting hyperplane, we mean that the hyperplane intersects S only at the point x.) Note that convexity of the set S is not required.

Proof. Suppose x is the minimum element of S, i.e., $x \preceq_K z$ for all $z \in S$, and let $\lambda \succ_{K^*} 0$. Let $z \in S, z \neq x$. Since x is the minimum element of S, we have $z - x \succeq_K 0$. From $\lambda \succ_{K^*} 0$ and $z - x \succeq_K 0, z - x \neq 0$, we conclude $\lambda^T(z - x) > 0$. Since z is an arbitrary element of S, not equal to x, this shows that x is the unique minimizer of $\lambda^T z$ over $z \in S$. Conversely, suppose that for all $\lambda \succ_{K^*} 0, x$ is the unique minimizer of $\lambda^T z$ over $z \in S$, but x is not the minimum element of S. Then there exists $z \in S$ with $z \not\succeq Kx$. Since $z - x \not\succeq_K 0$, there exists $\tilde{\lambda} \succeq_{K^*} 0$ with $\tilde{\lambda}^T(z - x) < 0$. Hence $\lambda^T(z - x) < 0$ for $\lambda \succ_{K^*} 0$ in the neighborhood of $\tilde{\lambda}$. This contradicts the assumption that x is the unique minimizer of $\lambda^T z$ over S. ■

Theorem 1.2.13 If $\lambda \succ_{K^*} 0$ and x minimizes $\lambda^T z$ over $z \in S$, then x is minimal.

Proof. To show this, suppose that $\lambda \succ_{K^*} 0$, and x minimizes $\lambda^T z$ over S, but x is not minimal, i.e., there exists a $z \in S, z \neq x$, and $z \preceq_K x$. Then $\lambda^T(x - z) > 0$, which contradicts our assumption that x is the minimizer of $\lambda^T z$ over S. ■

R The converse is in general false: a point x can be minimal in S, but not a minimizer of $\lambda^T z$ over $z \in S$, for any λ.

Theorem 1.2.14 Provided the set S is convex, we can say that for any minimal element x there exists a nonzero $\lambda \succeq_{K^*} 0$ such that x minimizes $\lambda^T z$ over $z \in S$.

Proof. To show this, suppose x is minimal, which means that $((x-K)\backslash\{x\}) \cap S = \emptyset$. Applying the separating hyperplane theorem to the convex sets $(x-K)\backslash\{x\}$ and S, we conclude that there is a $\lambda \neq 0$ and μ such that $\lambda^T(x-y) \leq \mu$ for all $y \in K$ and $\lambda^T z \geq \mu$ for all $z \in S$. From the first inequality we conclude $\lambda \succeq_{K^*} 0$. Since $x \in S$ and $x \in x-K$, we have $\lambda^T x = \mu$, so the second inequality implies that μ is the minimum value of $\lambda^T z$ over S. Therefore, x is a minimizer of $\lambda^T z$ over S, where $\lambda \neq 0, \lambda \succeq_{K^*} 0$ ■

R This converse theorem cannot be strengthened to $\lambda \succ_{K^*} 0$. Examples show that a point x can be a minimal point of a convex set S, but not a minimizer of $\lambda^T z$ over $z \in S$ for any $\lambda \succ_{K^*} 0$. Nor is it true that any minimizer of $\lambda^T z$ over $z \in S$, with $\lambda \succeq_{K^*} 0$, is minimal.

1.3 Convex function

Definition 1.3.1 — Convex function. A function $f : \mathbf{R}^n \to \mathbf{R}$ is **convex** if dom f is a convex set and if for all $x, y \in \operatorname{dom} f$, and θ with $0 \leq \theta \leq 1$, we have

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y) \tag{1.9}$$

Proposition 1.3.1 — Properties of Convex function.

1. Geometrically, Chord over function: this inequality means that the line segment between $(x, f(x))$ and $(y, f(y))$, which is the chord from x to y, lies above the graph of f.
2. A function f is **strictly convex** if strict inequality holds whenever $x \neq y$ and $0 < \theta < 1$.
3. We say f is **concave** if $-f$ is convex, and strictly concave if $-f$ is strictly convex.
4. All affine (and therefore also linear) functions are both convex and concave. Conversely, any function that is convex and concave is affine.
5. A function is convex if and only if it is convex when restricted to any line that intersects its domain (Just need to consider the situation of R^1). In other words f is convex if and only if for all $x \in$ dom f and all v, the function $g(t) = f(x+tv)$ is convex (on its domain , $\{t \mid x + tv \in \operatorname{dom} f\}$). This property is very useful, since it allows us to check whether a function is convex by restricting it to a line.
6. Continuity of convex function: A convex function is continuous on the relative interior of its domain; it can have discontinuities only on its relative boundary.

1.3.1 First order condition

Theorem 1.3.2 Suppose f is differentiable (*i.e.*, its gradient ∇f exists at each point in $\operatorname{dom} f$, which is open). Then f is **convex** if and only if $\operatorname{dom} f$ is convex and

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) \tag{1.10}$$

holds for all $x, y \in \operatorname{dom} f$. The affine function of y given by $f(x) + \nabla f(x)^T (y - x)$ is, of course, the first-order Taylor approximation of f near x.

R

1. The inequality (1.10) states that for a convex function, the first-order Taylor approximation is in fact a global underestimator of the function. Conversely, if the first-order Taylor approximation of a function is always a global underestimator of the function, then the function is convex.
2. The inequality (1.10) shows that from local information about a convex function (i.e., its value and derivative at a point) we can derive global information (*i.e.*, a global underestimator of it). As one simple example, the inequality (1.10) shows that if $\nabla f(x) = 0$, then for all $y \in \operatorname{dom} f, f(y) \geq f(x)$, i.e., x is a global minimizer of the function f.

1. Strict convexity can also be characterized by a first-order condition: f is strictly convex if and only if $\operatorname{dom} f$ is convex and for $x, y \in \operatorname{dom} f, x \neq y$, we have

$$f(y) > f(x) + \nabla f(x)^T (y - x)$$

2. For concave functions we have the corresponding characterization: f is concave if and only if $\operatorname{dom} f$ is convex and

$$f(y) \leq f(x) + \nabla f(x)^T (y - x)$$

for all $x, y \in \operatorname{dom} f$

Analysis 1.1 We just need to prove the situation of R^1, and using the $g(t) = f(x + tv)$ ∎

Proof. To prove (1.10), we first consider the case $n = 1$: We show that a differentiable function $f : \mathbf{R} \to \mathbf{R}$ is convex if and only if

$$f(y) \geq f(x) + f'(x)(y - x) \tag{1.11}$$

for all x and y in $\operatorname{dom} f$. Assume first that f is convex and $x, y \in \operatorname{dom} f$. Since $\operatorname{dom} f$ is convex (*i.e.*, an interval), we conclude that for all $0 < t \leq 1, x + t(y - x) \in \operatorname{dom} f$, and by convexity of f

$$f(x + t(y - x)) \leq (1 - t) f(x) + t f(y)$$

If we divide both sides by t, we obtain

$$f(y) \geq f(x) + \frac{f(x+t(y-x)) - f(x)}{t \cdot (y-x)} \cdot (y-x)$$

and taking the limit as $t \to 0$ yields (1.11).

To show sufficiency, assume the function satisfies (1.11) for all x and y in $\operatorname{dom} f$ (which is an interval). Choose any $x \neq y$, and $0 \leq \theta \leq 1$, and let $z = \theta x + (1-\theta)y$. Applying (1.11) twice yields

$$f(x) \geq f(z) + f'(z)(x-z), \quad f(y) \geq f(z) + f'(z)(y-z)$$

Multiplying the first inequality by θ, the second by $1-\theta$, and adding them yields

$$\theta f(x) + (1-\theta) f(y) \geq f(z)$$

which proves that f is convex.

Now we can prove the general case, with $f : \mathbf{R}^n \to \mathbf{R}$. Let $x, y \in \mathbf{R}^n$ and consider f restricted to the line passing through them, *i.e.*, the function defined by $g(t) = f(ty + (1-t)x)$, so $g'(t) = \nabla f(ty + (1-t)x)^T (y-x)$. First assume f is convex, which implies g is convex, so by the argument above we have $g(1) \geq g(0) + g'(0)$, which means

$$f(y) \geq f(x) + \nabla f(x)^T (y-x)$$

Now assume that this inequality holds for any x and y, so if $ty + (1-t)x \in \operatorname{dom} f$ and $\tilde{t}y + (1-\tilde{t})x \in \operatorname{dom} f$, we have

$$f(ty + (1-t)x) \geq f(\tilde{t}y + (1-\tilde{t})x) + \nabla f(\tilde{t}y + (1-\tilde{t})x)^T (y-x)(t-\tilde{t})$$

i.e., $g(t) \geq g(\tilde{t}) + g'(\tilde{t})(t-\tilde{t})$. We have seen that this implies that g is convex. ■

1.3.2 Second-order conditions

Theorem 1.3.3 — Second-order conditions. We now assume that f is twice differentiable, that is, its Hessian or second derivative $\nabla^2 f$ exists at each point in $\operatorname{dom} f$, which is open. Then f is **convex** if and only if $\operatorname{dom} f$ is convex and its Hessian is positive semidefinite: for all $x \in \operatorname{dom} f$,

$$\nabla^2 f(x) \succeq 0$$

For a function on $\mathbf{R}$, this reduces to the simple condition $f''(x) \geq 0$ (and $\operatorname{dom} f$ convex, i.e., an interval), which means that the derivative (f') is nondecreasing. The condition $\nabla^2 f(x) \succeq 0$ can be interpreted geometrically as the requirement that the graph of the function have positive (upward) curvature at x.

1. Similarly, f is **concave** if and only if $\operatorname{dom} f$ is convex and $\nabla^2 f(x) \preceq 0$ for all $x \in \operatorname{dom} f$.
2. If $\nabla^2 f(x) \succ 0$ for all $x \in \operatorname{dom} f$, then f is strictly convex.

3. The converse of strictly convex condition, however, is not true: for example, the function $f : \mathbf{R} \to \mathbf{R}$ given by $f(x) = x^4$ is strictly convex but has zero second derivative at $x = 0$.

Proposition 1.3.4 1. Quadratic functions. Consider the quadratic function $f : \mathbf{R}^n \to \mathbf{R}$, with $\operatorname{dom} f = \mathbf{R}^n$, given by

$$f(x) = (1/2)x^T Px + q^T x + r$$

with $P \in \mathbf{S}^n, q \in \mathbf{R}^n$, and $r \in \mathbf{R}$. Since $\nabla^2 f(x) = P$ for all x, f is convex if and only if $P \succeq 0$ (and concave if and only if $P \preceq 0$).

2. The separate requirement that dom f be convex cannot be dropped. For example, the function $f(x) = 1/x^2$, with $\operatorname{dom} f = \{x \in \mathbf{R} \mid x \neq 0\}$, satisfies $f''(x) > 0$ for all $x \in \operatorname{dom} f$, but is not a convex function.

■ **Example 1.8** 1. Exponential. e^{ax} is convex on $\mathbf{R}$, for any $a \in \mathbf{R}$

2. Powers. x^a is convex on $\mathbf{R}_{++}$ when $a \geq 1$ or $a \leq 0$, and concave for $0 \leq a \leq 1$
3. Powers of absolute value. $|x|^p$, for $p \geq 1$, is convex on $\mathbf{R}$.
4. Logarithm. $\log x$ is concave on $\mathbf{R}_{++}$
5. Negative entropy. $x \log x$ (either on $\mathbf{R}_{++}$, or on $\mathbf{R}_{+}$, defined as 0 for $x = 0$) is convex.
6. Norms. Every norm on $\mathbf{R}^n$ is convex. (triangle inequality, and homogeneity.)
7. Max function. $f(x) = \max\{x_1, \ldots, x_n\}$ is convex on $\mathbf{R}^n$
8. Quadratic-over-linear function. The function $f(x,y) = x^2/y$, with

$$\operatorname{dom} f = \mathbf{R} \times \mathbf{R}_{++} = \left\{(x,y) \in \mathbf{R}^2 \mid y > 0\right\}$$

is convex.

9. Log-sum-exp. The function $f(x) = \log\left(e^{x_1} + \cdots + e^{x_n}\right)$ is convex on $\mathbf{R}^n$. This function can be interpreted as a differentiable (in fact, analytic) approximation of the max function, since

$$\max\{x_1, \ldots, x_n\} \leq f(x) \leq \max\{x_1, \ldots, x_n\} + \log n$$

for all x. (The second inequality is tight when all components of x are equal.)

10. Geometric mean. The geometric mean $f(x) = \left(\prod_{i=1}^n x_i\right)^{1/n}$ is concave on $\operatorname{dom} f = \mathbf{R}^n_{++}$
11. Log-determinant. The function $f(X) = \log\det X$ is concave on $\operatorname{dom} f = \mathbf{S}^n_{++}$

Proof. 1. Log-sum-exp. The Hessian of the log-sum-exp function is

$$\nabla^2 f(x) = \frac{1}{(1^T z)^2}\left(\left(1^T z\right)\operatorname{diag}(z) - zz^T\right)$$

where $z = (e^{x_1}, \dots, e^{x_n})$. To verify that $\nabla^2 f(x) \succeq 0$ we must show that for all v, $v^T \nabla^2 f(x) v \geq 0$, i.e.

$$v^T \nabla^2 f(x) v = \frac{1}{(1^T z)^2} \left(\left(\sum_{i=1}^n z_i \right) \left(\sum_{i=1}^n v_i^2 z_i \right) - \left(\sum_{i=1}^n v_i z_i \right)^2 \right) \geq 0$$

But this follows from the Cauchy-Schwarz inequality $\left(a^T a\right)\left(b^T b\right) \geq \left(a^T b\right)^2$ applied to the vectors with components $a_i = v_i \sqrt{z_i}, b_i = \sqrt{z_i}$

2. Geometric mean $f(x) = \left(\prod_{i=1}^n x_i\right)^{1/n}$ is concave on $\operatorname{dom} f = \mathbf{R}_{++}^n$. Its Hessian $\nabla^2 f(x)$ is given by

$$\frac{\partial^2 f(x)}{\partial x_k^2} = -(n-1) \frac{\left(\prod_{i=1}^n x_i\right)^{1/n}}{n^2 x_k^2}, \quad \frac{\partial^2 f(x)}{\partial x_k \partial x_l} = \frac{\left(\prod_{i=1}^n x_i\right)^{1/n}}{n^2 x_k x_l} \quad \text{for } k \neq l$$

and can be expressed as

$$\nabla^2 f(x) = -\frac{\prod_{i=1}^n x_i^{1/n}}{n^2} \left(n \operatorname{diag}\left(1/x_1^2, \dots, 1/x_n^2\right) - q q^T \right)$$

where $q_i = 1/x_i$. We must show that $\nabla^2 f(x) \preceq 0$, i.e., that

$$v^T \nabla^2 f(x) v = -\frac{\prod_{i=1}^n x_i^{1/n}}{n^2} \left(n \sum_{i=1}^n v_i^2 / x_i^2 - \left(\sum_{i=1}^n v_i / x_i \right)^2 \right) \leq 0$$

for all v. Again this follows from the Cauchy-Schwarz inequality $\left(a^T a\right)\left(b^T b\right) \geq \left(a^T b\right)^2$, applied to the vectors $a = \mathbf{1}$ and $b_i = v_i / x_i$

3. Log-determinant. For the function $f(X) = \log \det X$, we can verify concavity by considering an arbitrary line, given by $X = Z + tV$, where $Z, V \in \mathbf{S}^n$. We define $g(t) = f(Z + tV)$, and restrict g to the interval of values of t for which $Z + tV \succ 0$ Without loss of generality, we can assume that $t = 0$ is inside this interval, i.e., $Z \succ 0$. We have

$$\begin{aligned} g(t) &= \log \det(Z + tV) \\ &= \log \det \left(Z^{1/2} \left(I + t Z^{-1/2} V Z^{-1/2} \right) Z^{1/2} \right) \\ &= \sum_{i=1}^n \log\left(1 + t\lambda_i\right) + \log \det Z \end{aligned}$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of $Z^{-1/2} V Z^{-1/2}$. Therefore we have

$$g'(t) = \sum_{i=1}^n \frac{\lambda_i}{1 + t\lambda_i}, \quad g''(t) = -\sum_{i=1}^n \frac{\lambda_i^2}{(1 + t\lambda_i)^2}$$

since $g''(t) \leq 0$, we conclude that f is concave.

■

Definition 1.3.2 — establish convexity of a set–α -sublevel set. The α -sublevel set of a function $f : \mathbf{R}^n \to \mathbf{R}$ is defined as

$$C_\alpha = \{x \in \operatorname{dom} f \mid f(x) \le \alpha\}$$

Proposition 1.3.5

1. Sublevel sets of a convex function are convex, for any value of α. (definition)
2. The converse is not true: a function can have all its sublevel sets convex, but not be a convex function. For example, $f(x) = -e^x$ is not convex on $\mathbf{R}$ (indeed, it is strictly concave) but all its sublevel sets are convex.
3. If f is concave, then its α -superlevel set, given by $\{x \in \operatorname{dom} f \mid f(x) \ge \alpha\}$, is a convex set.

■ **Example 1.9** The geometric and arithmetic means of $x \in \mathbf{R}^n_+$ are, respectively,

$$G(x) = \left(\prod_{i=1}^n x_i\right)^{1/n}, \quad A(x) = \frac{1}{n}\sum_{i=1}^n x_i$$

(where we take $0^{1/n} = 0$ in our definition of G). The arithmetic-geometric mean inequality states that $G(x) \le A(x)$. Suppose $0 \le \alpha \le 1$, and consider the set

$$\{x \in \mathbf{R}^n_+ \mid G(x) \ge \alpha A(x)\}$$

i.e., the set of vectors with geometric mean at least as large as a factor α times the arithmetic mean. This set is convex, since it is the 0 -superlevel set of the function $G(x) - \alpha A(x)$, which is concave. In fact, the set is positively homogeneous, so it is a convex cone.

Definition 1.3.3 — Epigraph. The graph of a function $f : \mathbf{R}^n \to \mathbf{R}$ is defined as

$$\{(x, f(x)) \mid x \in \operatorname{dom} f\}$$

which is a subset of $\mathbf{R}^{n+1}$. The **epigraph** of a function $f : \mathbf{R}^n \to \mathbf{R}$ is defined as

$$\operatorname{epi} f = \{(x,t) \mid x \in \operatorname{dom} f, f(x) \le t\}$$

which is a subset of $\mathbf{R}^{n+1}$. ('Epi' means 'above' so epigraph means 'above the graph'.)

The link between convex sets and convex functions is via the epigraph: A function is convex if and only if its epigraph is a convex set. A function is concave if and only if its hypograph, defined as

$$\operatorname{hypo} f = \{(x,t) \mid t \le f(x)\}$$

is a convex set.

■ **Example 1.10 — Matrix fractional function.** The function $f : \mathbf{R}^n \times \mathbf{S}^n \to \mathbf{R}$, defined as

$$f(x,Y) = x^T Y^{-1} x$$

is convex on $\operatorname{dom} f = \mathbf{R}^n \times \mathbf{S}^n_{++}$. (This generalizes the quadratic-over-linear function $f(x,y) = x^2/y$, with $\operatorname{dom} f = \mathbf{R} \times \mathbf{R}_{++}$). One easy way to establish convexity of f is via its epigraph:

$$\begin{aligned}\mathbf{epi} f &= \left\{(x,Y,t) \mid Y \succ 0, x^T Y^{-1} x \le t\right\} \\ &= \left\{(x,Y,t) \mid \begin{bmatrix} Y & x \\ x^T & t \end{bmatrix} \succeq 0, Y \succ 0\right\}\end{aligned}$$

using the Schur complement condition for positive semidefiniteness of a block matrix (see §A.5.5). The last condition is a linear matrix inequality in (x,Y,t), and therefore epi f is convex.

For the special case $n=1$, the matrix fractional function reduces to the quadraticover-linear function x^2/y, and the associated LMI representation is

$$\begin{bmatrix} y & x \\ x & t \end{bmatrix} \succeq 0, \quad y > 0$$

■ **Example 1.11** Many results for convex functions can be proved (or interpreted) geometrically using epigraphs, and applying results for convex sets. As an example, consider the first-order condition for convexity:

$$f(y) \ge f(x) + \nabla f(x)^T (y-x)$$

where f is convex and $x,y \in \operatorname{dom} f$. We can interpret this basic inequality geometrically in terms of epi f. If $(y,t) \in$ epi f, then

$$t \ge f(y) \ge f(x) + \nabla f(x)^T (y-x)$$

For a differentiable convex function f, the vector $(\nabla f(x), -1)$ defines a supporting hyperplane to the epigraph of f at x.

We can express this as:

$$(y,t) \in \mathbf{epi} f \Longrightarrow \begin{bmatrix} \nabla f(x) \\ -1 \end{bmatrix}^T \left(\begin{bmatrix} y \\ t \end{bmatrix} - \begin{bmatrix} x \\ f(x) \end{bmatrix} \right) \le 0$$

This means that the hyperplane defined by $(\nabla f(x), -1)$ supports epi f at the boundary point $(x, f(x))$.

1.3.3 Jensen's inequality and extensions

Definition 1.3.4 1. The basic inequality

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$$

is sometimes called **Jensen's inequality**.

2. More than two points: If f is convex, $x_1, \ldots, x_k \in \operatorname{dom} f$, and $\theta_1, \ldots, \theta_k \geq 0$ with $\theta_1 + \cdots + \theta_k = 1$, then

$$f(\theta_1 x_1 + \cdots + \theta_k x_k) \leq \theta_1 f(x_1) + \cdots + \theta_k f(x_k)$$

3. Extend to infinite sums, integrals, and expected values. If $p(x) \geq 0$ on $S \subseteq \operatorname{dom} f, \int_S p(x)dx = 1$, then

$$f\left(\int_S p(x)x dx\right) \leq \int_S f(x)p(x)dx$$

provided the integrals exist.

4. Probability measure with support in dom f. If x is a random variable such that $x \in \operatorname{dom} f$ with probability one, and f is convex, then we have

$$f(\mathbf{E}x) \leq \mathbf{E}f(x) \tag{1.12}$$

provided the expectations exist.

5. Taking the random variable x to have support $\{x_1, x_2\}$, with prob $(x = x_1) = \theta$, prob $(x = x_2) = 1 - \theta$. Thus the inequality (1.12) characterizes convexity: If f is not convex, there is a random variable x, with $x \in \operatorname{dom} f$ with probability one, such that $f(\mathbf{E}x) > \mathbf{E}f(x)$.

6. Interpret (1.12) as follows. Suppose $x \in \operatorname{dom} f \subseteq \mathbf{R}^n$ and z is any zero mean random vector in $\mathbf{R}^n$. Then we have

$$\mathbf{E}f(x+z) \geq f(x)$$

Thus, randomization or dithering (*i.e.*, adding a zero mean random vector to the argument) cannot decrease the value of a convex function on average.

7.

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2}$$

1.3.4 Inequalities

Theorem 1.3.6 — Arithmetic-Geometric mean inequality. arithmetic-geometric mean inequality:

$$\sqrt{ab} \leq (a+b)/2 \tag{1.13}$$

for $a,b \geq 0$. The function $-\log x$ is convex; Jensen's inequality with $\theta = 1/2$ yields

$$-\log\left(\frac{a+b}{2}\right) \leq \frac{-\log a - \log b}{2}$$

Taking the exponential of both sides yields (1.13).

Theorem 1.3.7 — Hölder's inequality. Hölder's inequality: for $p > 1, 1/p + 1/q = 1$, and $x, y \in \mathbf{R}^n$

$$\sum_{i=1}^{n} x_i y_i \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \left(\sum_{i=1}^{n} |y_i|^q\right)^{1/q}$$

By convexity of $-\log x$, and Jensen's inequality with general θ, we obtain the more general arithmetic-geometric mean inequality

$$a^\theta b^{1-\theta} \leq \theta a + (1-\theta) b$$

valid for $a, b \geq 0$ and $0 \leq \theta \leq 1$. Applying this with

$$a = \frac{|x_i|^p}{\sum_{j=1}^n |x_j|^p}, \quad b = \frac{|y_i|^q}{\sum_{j=1}^n |y_j|^q}, \quad \theta = 1/p$$

yields

$$\left(\frac{|x_i|^p}{\sum_{j=1}^n |x_j|^p}\right)^{1/p} \left(\frac{|y_i|^q}{\sum_{j=1}^n |y_j|^q}\right)^{1/q} \leq \frac{|x_i|^p}{p\sum_{j=1}^n |x_j|^p} + \frac{|y_i|^q}{q\sum_{j=1}^n |y_j|^q}$$

Summing over i then yields Hölder's inequality.

1.3.5 Operations that preserve convexity

Theorem 1.3.8 — Operations that preserve convexity.

1. A nonnegative weighted sum of convex functions,

$$f = w_1 f_1 + \cdots + w_m f_m$$

is convex.

2. Similarly, a nonnegative weighted sum of concave functions is concave. A nonnegative, nonzero weighted sum of strictly convex (concave) functions is strictly convex (concave).

3. These properties extend to infinite sums and integrals. For example if $f(x,y)$ is convex in x for each $y \in \mathcal{A}$, and $w(y) \geq 0$ for each $y \in \mathcal{A}$, then the function g defined as

$$g(x) = \int_{\mathcal{A}} w(y) f(x,y) dy$$

is convex in x (provided the integral exists).

4. The fact that convexity is preserved under nonnegative scaling and addition is easily verified directly, or can be seen in terms of the associated epigraphs. For example, if $w \geq 0$ and f is convex, we have

$$\mathrm{epi}(wf) = \begin{bmatrix} I & 0 \\ 0 & w \end{bmatrix} \mathrm{epi}\, f$$

which is convex because the image of a convex set under a linear mapping is convex.

5. Composition with an ane mapping: Suppose $f : \mathbf{R}^n \to \mathbf{R}, A \in \mathbf{R}^{n \times m}$, and $b \in \mathbf{R}^n$. Define $g : \mathbf{R}^m \to \mathbf{R}$ by

$$g(x) = f(Ax + b)$$

with $\mathrm{dom}\, g = \{x \mid Ax + b \in \mathrm{dom}\, f\}$. Then if f is convex, so is g; if f is concave, so is g

6. Pointwise maximun: If f_1 and f_2 are convex functions then their pointwise maximum f, defined by

$$f(x) = \max\{f_1(x), f_2(x)\}$$

with $\mathrm{dom}\, f = \mathrm{dom}\, f_1 \cap \mathrm{dom}\, f_2$, is also convex. This property is easily verified by definition.

7. It is easily shown that if $f_1, \ldots, f_m$ are convex, then their pointwise maximum

$$f(x) = \max\{f_1(x), \ldots, f_m(x)\}$$

is also convex.

8. Piecewise-linear functions:

$$f(x) = \max\left\{a_1^T x + b_1, \ldots, a_L^T x + b_L\right\}$$

defines a piecewise-linear (or really, affine) function (with L or fewer regions). It is convex since it is the pointwise maximum of affine functions.

9. The converse can also be shown: any piecewise-linear convex function with L or fewer regions can be expressed in this form.

■ **Example 1.12**

Sum of r largest components. For $x \in \mathbf{R}^n$ we denote by $x_{[i]}$ the i th largest component of x, i.e.,

$$x_{[1]} \geq x_{[2]} \geq \cdots \geq x_{[n]}$$

the sum of the r largest elements of x, is a convex function. This can be seen by writing it as

$$f(x) = \sum_{i=1}^{r} x_{[i]} = \max\{x_{i_1} + \cdots + x_{i_r} \mid 1 \leq i_1 < i_2 < \cdots < i_r \leq n\}$$

Since it is the pointwise maximum of $n!/(r!(n-r)!)$ linear functions, it is convex.

As an extension it can be shown that the function $\sum_{i=1}^{r} w_i x_{[i]}$ is convex, provided $w_1 \geq w_2 \geq \cdots \geq w_r \geq 0$.)

Theorem 1.3.9 1. Extends to the pointwise supremum over an infinite set of convex functions. If for each $y \in \mathcal{A}, f(x,y)$ is convex in x, then the function g, defined as

$$g(x) = \sup_{y \in \mathcal{A}} f(x,y)$$

is convex in x. Here the domain of g is

$$\operatorname{dom} g = \left\{ x \mid (x,y) \in \operatorname{dom} f \text{ for all } y \in \mathcal{A}, \sup_{y \in \mathcal{A}} f(x,y) < \infty \right\}$$

2. Similarly, the pointwise infimum of a set of concave functions is a concave function.
3. In terms of epigraphs, the pointwise supremum of functions corresponds to the intersection of epigraphs: with f, g, and $\mathcal{A}$ as defined previously, we have

$$\mathbf{epi} g = \bigcap_{y \in \mathcal{A}} \operatorname{epi} f(\cdot, y)$$

Thus, the result follows from the fact that the intersection of a family of convex sets is convex.

Expressing functions as the pointwise supremum of a family of ane functions.

■ **Example 1.13** 1. Support function of a set. Let $C \subseteq \mathbf{R}^n$, with $C \neq \emptyset$. The support function S_C associated with the set C is defined as

$$S_C(x) = \sup\left\{x^T y \mid y \in C\right\}$$

(and, naturally, $\operatorname{dom} S_C = \left\{x \mid \sup_{y \in C} x^T y < \infty\right\}$). For each $y \in C, x^T y$ is a linear function of x, so S_C is the pointwise supremum of a family of linear functions, hence convex.

2. Distance to farthest point of a set. Let $C \subseteq \mathbf{R}^n$. The distance (in any norm) to the farthest point of C,

$$f(x) = \sup_{y \in C} \|x - y\|$$

is convex. To see this, note that for any y, the function $\|x - y\|$ is convex in x. Since f is the pointwise supremum of a family of convex functions (indexed by $y \in C$), it is a convex function of x.

3. Maximum eigenvalue of a symmetric matrix. The function $f(X) = \lambda_{\max}(X)$, with $\operatorname{dom} f = \mathbf{S}^m$, is convex. To see this, we express f as

$$f(X) = \sup\left\{y^T X y \mid \|y\|_2 = 1\right\}$$

i.e., as the pointwise supremum of a family of linear functions of X (i.e., $y^T X y$) indexed by $y \in \mathbf{R}^m$

4. Norm of a matrix: Consider $f(X) = \|X\|_2$ with $\operatorname{dom} f = \mathbf{R}^{p \times q}$, where $\|\cdot\|_2$ denotes the spectral norm or maximum singular value. Convexity of f follows from

$$f(X) = \sup\left\{u^T X v \mid \|u\|_2 = 1, \|v\|_2 = 1\right\}$$

which shows it is the pointwise supremum of a family of linear functions of X.

As a generalization suppose $\|\cdot\|_a$ and $\|\cdot\|_b$ are norms on $\mathbf{R}^p$ and $\mathbf{R}^q$, respectively. The induced norm of a matrix $X \in \mathbf{R}^{p \times q}$ is defined as

$$\|X\|_{a,b} = \sup_{v \neq 0} \frac{\|Xv\|_a}{\|v\|_b}$$

(This reduces to the spectral norm when both norms are Euclidean.) The induced norm can be expressed as

$$\begin{aligned}\|X\|_{a,b} &= \sup\{\|Xv\|_a \mid \|v\|_b = 1\} \\ &= \sup\left\{u^T X v \mid \|u\|_{a*} = 1, \|v\|_b = 1\right\}\end{aligned}$$

where $\|\cdot\|_{a*}$ is the dual norm of $\|\cdot\|_a$, and we use the fact that

$$\|z\|_a = \sup\left\{u^T z \mid \|u\|_{a*} = 1\right\}$$

Since we have expressed $\|X\|_{a,b}$ as a supremum of linear functions of X, it is a convex function.

Theorem 1.3.10 — Representation as pointwise supremum of ane functions. Almost every convex function can be expressed as the pointwise supremum of a family of affine functions. For example, if $f : \mathbf{R}^n \to \mathbf{R}$ is convex, with $\operatorname{dom} f = \mathbf{R}^n$, then we have

$$f(x) = \sup\{g(x) \mid g \text{ affine}, g(z) \le f(z) \text{ for all } z\}$$

Proof. Suppose f is convex with $\operatorname{dom} f = \mathbf{R}^n$. The inequality

$$f(x) \ge \sup\{g(x) \mid g \text{ affine}, g(z) \le f(z) \text{ for all } z\}$$

is clear. To establish equality, we will show that for each $x \in \mathbf{R}^n$, there is an affine function g, which is a global underestimator of f, and satisfies $g(x) = f(x)$.

The epigraph of f is, of course, a convex set. Hence we can find a supporting hyperplane to it at $(x, f(x))$, i.e., $a \in \mathbf{R}^n$ and $b \in \mathbf{R}$ with $(a,b) \ne 0$ and

$$\begin{bmatrix} a \\ b \end{bmatrix}^T \begin{bmatrix} x-z \\ f(x)-t \end{bmatrix} \le 0$$

for all $(z,t) \in \mathbf{epi} f$. This means that

$$a^T(x-z) + b(f(x) - f(z) - s) \le 0 \tag{1.14}$$

for all $z \in \operatorname{dom} f = \mathbf{R}^n$ and all $s \ge 0$ (Since $(z,t) \in \operatorname{epi} f$ means $t = f(z) + s$ for some $s \ge 0$ and note that s is variable). For the inequality (1.14) to hold for all $s \ge 0$, we must have $b \ge 0$. If $b = 0$, then the inequality (1.14) reduces to $a^T(x-z) \le 0$ for all $z \in \mathbf{R}^n$, which implies $a = 0$ and contradicts $(a,b) \ne 0$. We conclude that $b > 0$, i.e., that the supporting hyperplane is not vertical. Using the fact that $b > 0$ we rewrite (1.14) for $s = 0$ as

$$g(z) = f(x) + (a/b)^T(x-z) \le f(z)$$

for all z. The function g is an affine underestimator of f, and satisfies $g(x) = f(x)$. ■

Composition

Definition 1.3.5 — Composition. In this section we examine conditions on $h : \mathbf{R}^k \to \mathbf{R}$ and $g : \mathbf{R}^n \to \mathbf{R}^k$ that guarantee convexity or concavity of their composition $f = h \circ g : \mathbf{R}^n \to \mathbf{R}$, defined by

$$f(x) = h(g(x)), \quad \operatorname{dom} f = \{x \in \operatorname{dom} g \mid g(x) \in \operatorname{dom} h\}$$

We first consider the case $k = 1$, so $h : \mathbf{R} \to \mathbf{R}$ and $g : \mathbf{R}^n \to \mathbf{R}$. We can restrict ourselves to the case $n = 1$ (by using g).

Assuming that h and g are twice differentiable, with $\operatorname{dom} g = \operatorname{dom} h = \mathbf{R}$. In this case, convexity of f reduces to $f'' \ge 0$ (meaning, $f''(x) \ge 0$ for all $x \in \mathbf{R}$)

The second derivative of the composition function $f = h \circ g$ is given by

$$f''(x) = h''(g(x))g'(x)^2 + h'(g(x))g''(x) \tag{1.15}$$

We have the following results:

1. f is convex if h is convex and nondecreasing, and g is convex,
2. f is convex if h is convex and nonincreasing, and g is concave,
3. f is concave if h is concave and nondecreasing, and g is concave,
4. f is concave if h is concave and nonincreasing, and g is convex.

In the general case $n > 1$, without assuming differentiability of h and g, or that $\operatorname{dom} g = \mathbf{R}^n$ and $\operatorname{dom} h = \mathbf{R}$:

1. f is convex if h is convex, $\tilde{h}$ is nondecreasing, and g is convex,
2. f is convex if h is convex, $\tilde{h}$ is nonincreasing, and g is concave,
3. f is concave if h is concave, $\tilde{h}$ is nondecreasing, and g is concave,
4. f is concave if h is concave, $\tilde{h}$ is nonincreasing, and g is convex.

Here $\tilde{h}$ denotes the extended-value extension of the function h, which assigns the value $\infty(-\infty)$ to points not in $\operatorname{dom} h$ for h convex (concave).

Theorem 1.3.11 Without assuming differentiability, we will prove the following composition theorem: if g is convex, h is convex, and $\tilde{h}$ is nondecreasing, then $f = h \circ g$ is convex.

Proof. Assume that $x, y \in \operatorname{dom} f$, and $0 \leq \theta \leq 1$. Since $x, y \in \operatorname{dom} f$, we have that $x, y \in \operatorname{dom} g$ and $g(x), g(y) \in \operatorname{dom} h$. Since dom g is convex, we conclude that $\theta x + (1-\theta)y \in \operatorname{dom} g$, and from convexity of g, we have

$$g(\theta x + (1-\theta)y) \leq \theta g(x) + (1-\theta)g(y) \tag{1.16}$$

Since $g(x), g(y) \in \operatorname{dom} h$, we conclude that $\theta g(x) + (1-\theta)g(y) \in \operatorname{dom} h$, i.e. the righthand side of (1.16) is in $\operatorname{dom} h$. Now we use the assumption that h is nondecreasing, which means that its domain extends infinitely in the negative direction. Since the righthand side of (1.16) is in $\operatorname{dom} h$, we conclude that the lefthand side, *i.e.*, $g(\theta x + (1-\theta)y) \in \operatorname{dom} h$. This means that $\theta x + (1-\theta)y \in \operatorname{dom} f$. At this point, we have shown that dom f is convex. Now using the fact that $\tilde{h}$ is nondecreasing and the inequality (1.16), we get

$$h(g(\theta x + (1-\theta)y)) \leq h(\theta g(x) + (1-\theta)g(y))$$

From convexity of h, we have

$$h(\theta g(x) + (1-\theta)g(y)) \leq \theta h(g(x)) + (1-\theta)h(g(y))$$

Putting two formula together, we have

$$h(g(\theta x + (1-\theta)y)) \leq \theta h(g(x)) + (1-\theta)h(g(y))$$

which proves the composition theorem. ■

■ **Example 1.14** The requirement that monotonicity hold for the extended-value extension $\tilde{h}$, and not just the function h, cannot be removed.

For example, consider the function $g(x) = x^2$, with $\operatorname{dom} g = \mathbf{R}$, and $h(x) = 0$, with $\operatorname{dom} h = [1,2]$. Here g is convex, and h is convex and nondecreasing. But the function $f = h \circ g$, given by

$$f(x) = 0, \quad \operatorname{dom} f = [-\sqrt{2}, -1] \cup [1, \sqrt{2}]$$

is not convex, since its domain is not convex. Here, of course, the function $\tilde{h}$ is *not* nondecreasing.

Vector composition–Case when $k \geq 1$

Suppose

$$f(x) = h(g(x)) = h(g_1(x), \ldots, g_k(x))$$

with $h : \mathbf{R}^k \to \mathbf{R}, g_i : \mathbf{R}^n \to \mathbf{R}$. Again without loss of generality we can assume n =1. Similar to the previous case, We have

$$f''(x) = g'(x)^T \nabla^2 h(g(x)) g'(x) + \nabla h(g(x))^T g''(x)$$

We can derive many rules, for example:

1. f is convex if h is convex, h is nondecreasing in each argument, and g_i are convex,
2. f is convex if h is convex, h is nonincreasing in each argument, and g_i are concave,
3. f is concave if h is concave, h is nondecreasing in each argument, and g_i are concave.

Similar composition results hold in general, with $n > 1$, no assumption of differentiability of h or g, and general domains. For the general results, the monotonicity condition on h must hold for the extended-value extension $\tilde{h}$.

To understand the meaning of the condition that the extended-value extension $\tilde{h}$ be monotonic, we consider the case where $h : \mathbf{R}^k \to \mathbf{R}$ is convex, and $\tilde{h}$ nondecreasing, i.e., whenever $u \preceq v$, we have $\tilde{h}(u) \leq \tilde{h}(v)$. This implies that if $v \in \operatorname{dom} h$, then so is u : the domain of h must extend infinitely in the $-\mathbf{R}_+^k$ directions. We can express this compactly as $\operatorname{dom} h - \mathbf{R}_+^k = \operatorname{dom} h$

■ **Example 1.15**

1. Let $h(z) = z_{[1]} + \cdots + z_{[r]}$, the sum of the r largest components of $z \in \mathbf{R}^k$. Then h is convex and nondecreasing in each argument. Suppose $g_1, \ldots, g_k$ are convex functions on $\mathbf{R}^n$. Then the composition function $f = h \circ g$, i.e., the pointwise sum of the r largest g_i 's, is convex.
2. The function $h(z) = \log\left(\sum_{i=1}^k e^{z_i}\right)$ is convex and nondecreasing in each argument, so $\log\left(\sum_{i=1}^k e^{g_i}\right)$ is convex whenever g_i are.
3. For $0 < p \leq 1$, the function $h(z) = \left(\sum_{i=1}^k z_i^p\right)^{1/p}$ on $\mathbf{R}_+^k$ is concave, and its extension (which has the value $-\infty$ for $z \not\succeq 0$) is nondecreasing in each component. So if g_i are concave and nonnegative, we conclude that $f(x) = \left(\sum_{i=1}^k g_i(x)^p\right)^{1/p}$ is concave.

4. Suppose $p \geq 1$, and $g_1, \ldots, g_k$ are convex and nonnegative. Then the function $\left(\sum_{i=1}^k g_i(x)^p\right)^{1/p}$ is convex.

   To show this, we consider the function $h : \mathbf{R}^k \to \mathbf{R}$ defined as
$$h(z) = \left(\sum_{i=1}^k \max\{z_i, 0\}^p\right)^{1/p}$$
   with $\operatorname{dom} h = \mathbf{R}^k$, so $h = \tilde{h}$. This function is convex, and nondecreasing, so we conclude $h(g(x))$ is a convex function of x. For $z \succeq 0$, we have $h(z) = \left(\sum_{i=1}^k z_i^p\right)^{1/p}$, so our conclusion is that $\left(\sum_{i=1}^k g_i(x)^p\right)^{1/p}$ is convex.

5. The geometric mean $h(z) = \left(\prod_{i=1}^k z_i\right)^{1/k}$ on $\mathbf{R}_+^k$ is concave and its extension is nondecreasing in each argument. It follows that if $g_1, \ldots, g_k$ are nonnegative concave functions, then so is their geometric mean, $\left(\prod_{i=1}^k g_i\right)^{1/k}$

Minimization

If f is convex in (x, y), and C is a convex nonempty set, then the function
$$g(x) = \inf_{y \in C} f(x, y) \tag{1.17}$$
is convex in x, provided $g(x) > -\infty$ for all x. The domain of g is the projection of dom f on its x -coordinates, i.e.,
$$\operatorname{dom} g = \{x \mid (x, y) \in \operatorname{dom} f \text{ for some } y \in C\}$$

Proof. We prove this by verifying Jensen's inequality for $x_1, x_2 \in \operatorname{dom} g$. Let $\epsilon > 0$. Then there are $y_1, y_2 \in C$ such that $f(x_i, y_i) \leq g(x_i) + \epsilon$ for $i = 1, 2$. Now let $\theta \in [0, 1]$. We have
$$\begin{aligned} g(\theta x_1 + (1-\theta)x_2) &= \inf_{y \in C} f(\theta x_1 + (1-\theta)x_2, y) \\ &\leq f(\theta x_1 + (1-\theta)x_2, \theta_1 + (1-\theta)y_2) \\ &\leq \theta f(x_1, y_1) + (1-\theta) f(x_2, y_2) \\ &\leq \theta g(x_1) + (1-\theta) g(x_2) + \epsilon \end{aligned}$$
Since this holds for any $\epsilon > 0$, we have
$$g(\theta x_1 + (1-\theta)x_2) \leq \theta g(x_1) + (1-\theta) g(x_2)$$
■

The result can also be seen in terms of epigraphs. With f, g, and C defined as in (1.17), and assuming the infimum over $y \in C$ is attained for each x, we have
$$\operatorname{epi} g = \{(x, t) \mid (x, y, t) \in \operatorname{epi} f \text{ for some } y \in C\}$$
Thus epi g is convex, since it is the projection of a convex set on some of its components.

■ **Example 1.16** 1. Schur complement. Suppose the quadratic function

$$f(x,y) = x^T Ax + 2x^T By + y^T Cy$$

(where A and C are symmetric) is convex in (x,y), which means

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succeq 0$$

We can express $g(x) = \inf_y f(x,y)$ as

$$g(x) = x^T \left(A - BC^{\dagger} B^T \right) x$$

where $C^{\dagger}$ is the pseudo-inverse of C (see $.A.5.4). By the minimization rule, g is convex, so we conclude that $A - BC^{\dagger}B^T \succeq 0$. If C is invertible, *i.e.*,$C \succ 0$, then the matrix $A - BC^{-1}B^T$ is called the **Schur complement** of C in the matrix

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$$

2. Distance to a set. The distance of a point x to a set $S \subseteq \mathbf{R}^n$, in the norm $\|\cdot\|$, is defined as

$$\operatorname{dist}(x,S) = \inf_{y \in S} \|x - y\|$$

The function $\|x - y\|$ is convex in (x,y), so if the set S is convex, the distance function $\operatorname{dist}(x,S)$ is a convex function of x.

3. Suppose h is convex. Then the function g defined as

$$g(x) = \inf\{h(y) \mid Ay = x\}$$

is convex. To see this, we define f by

$$f(x,y) = \begin{cases} h(y) & \text{if } Ay = x \\ \infty & \text{otherwise} \end{cases}$$

which is convex in (x,y). Then g is the minimum of f over y, and hence is convex.

Perspective of a function

Definition 1.3.6 — Perspective. If $f : \mathbf{R}^n \to \mathbf{R}$, then the **perspective** of f is the function $g : \mathbf{R}^{n+1} \to \mathbf{R}$ defined by

$$g(x,t) = tf(x/t)$$

with domain

$$\operatorname{dom} g = \{(x,t) \mid x/t \in \operatorname{dom} f, t > 0\}$$

The perspective operation preserves convexity: If f is a convex function, then so is its perspective function g. Similarly, if f is concave, then so is g.

Proof. For $t > 0$ we have

$$\begin{aligned}(x,t,s) \in \mathbf{epi} g &\Longleftrightarrow tf(x/t) \leq s \\ &\Longleftrightarrow f(x/t) \leq s/t \\ &\Longleftrightarrow (x/t, s/t) \in \mathbf{epi} f\end{aligned}$$

Therefore epi g is the inverse image of epi f under the perspective mapping that takes (u,v,w) to $(u,w)/v$. If epi g is convex, so the function g is convex. ■

■ **Example 1.17 — Negative logarithm.** Consider the convex function $f(x) = -\log x$ on $\mathbf{R}_{++}$. Its perspective is

$$g(x,t) = -t\log(x/t) = t\log(t/x) = t\log t - t\log x$$

and is convex on $\mathbf{R}^2_{++}$. The function g is called the **relative entropy** of t and x. For $x = 1, g$ reduces to the negative entropy function.

From convexity of g we can establish convexity or concavity of several interesting related functions. First, the relative entropy of two vectors $u, v \in \mathbf{R}^n_{++}$, defined as

$$\sum_{i=1}^{n} u_i \log(u_i/v_i)$$

is convex in (u,v), since it is a sum of relative entropies of u_i, v_i.

A closely related function is the **Kullback-Leibler divergence** between $u, v \in \mathbf{R}^n_{++}$, given by

$$D_{\mathrm{kl}}(u,v) = \sum_{i=1}^{n} (u_i \log(u_i/v_i) - u_i + v_i)$$

which is convex, since it is the relative entropy plus a linear function of (u,v). The Kullback-Leibler divergence satisfies $D_{\mathrm{kl}}(u,v) \geq 0$, and $D_{\mathrm{kl}}(u,v) = 0$ if and only if $u = v$, and so can be used as a measure of deviation between two positive vectors. (Note that the relative entropy and the Kullback-Leibler divergence are the same when u and v are probability vectors, i.e., satisfy $\mathbf{1}^T u = \mathbf{1}^T v = 1$.)

If we take $v_i = \mathbf{1}^T u$ in the relative entropy function, we obtain the concave (and homogeneous) function of $u \in \mathbf{R}^n_{++}$ given by

$$\sum_{i=1}^{n} u_i \log\left(\mathbf{1}^T u/u_i\right) = \left(\mathbf{1}^T u\right) \sum_{i=1}^{n} z_i \log(1/z_i)$$

where $z = u/\left(\mathbf{1}^T u\right)$, which is called the **normalized entropy function**. The vector $z = u/\mathbf{1}^T u$ is a normalized vector or probability distribution, since its components sum to one; the normalized entropy of u is $\mathbf{1}^T u$ times the entropy of this normalized distribution.

1.3.6 The conjugate function

Definition 1.3.7 — Conjugate. Let $f : \mathbf{R}^n \to \mathbf{R}$. The function $f^* : \mathbf{R}^n \to \mathbf{R}$, defined as

$$f^*(y) = \sup_{x \in \operatorname{dom} f} \left(y^T x - f(x) \right)$$

is called the **conjugate** of the function f.

We see immediately that f^* is a convex function, since it is the pointwise supremum of a family of convex (indeed, affine) functions of y (Conjugate is the function of y, but the sup is obtained from the dom f of x). This is true whether or not f is convex. (Note that when f is convex, the subscript $x \in \operatorname{dom} f$ is not necessary since, by convention, $y^T x - f(x) = -\infty$ for $x \notin \operatorname{dom} f$.)

■ **Example 1.18 — log-determinant.** We consider $f(X) = \log\det X^{-1}$ on $\mathbf{S}^n_{++}$. The conjugate function is defined as

$$f^*(Y) = \sup_{X \succ 0} (\operatorname{tr}(YX) + \log\det X)$$

since $\operatorname{tr}(YX)$ is the standard inner product on $\mathbf{S}^n$. We first show that $\operatorname{tr}(YX)+$ log det X is unbounded above unless $Y \prec 0$. If $Y \not\prec 0$, then Y has an eigenvector v, with $\|v\|_2 = 1$, and eigenvalue $\lambda \geq 0$. Taking $X = I + tvv^T$ we find that

$$\operatorname{tr}(YX) + \log\det X = \operatorname{tr} Y + t\lambda + \log\det\left(I + tvv^T\right) = \operatorname{tr} Y + t\lambda + \log(1+t)$$

which is unbounded above as $t \to \infty$. Now consider the case $Y \prec 0$. We can find the maximizing X by setting the gradient with respect to X equal to zero:

$$\nabla_X(\operatorname{tr}(YX) + \log\det X) = Y + X^{-1} = 0$$

which yields $X = -Y^{-1}$ (which is, indeed, positive definite). Therefore we have

$$f^*(Y) = \log\det(-Y)^{-1} - n$$

with $\operatorname{dom} f^* = -\mathbf{S}^n_{++}$

■ **Example 1.19 — Indicator function.** Let I_S be the indicator function of a (not necessarily convex) set $S \subseteq \mathbf{R}^n$, i.e., $I_S(x) = 0$ on $\operatorname{dom} I_S = S$. Its conjugate is

$$I_S^*(y) = \sup_{x \in S} y^T x$$

which is the **support function** of the set S.

■ **Example 1.20 — log -sum-exp function.** To derive the conjugate of the log-sum-exp function $f(x) = \log\left(\sum_{i=1}^n e^{x_i}\right)$, we first determine the values of y for which the maximum over x of $y^T x - f(x)$ is attained. By setting the gradient with respect to x equal to zero, we obtain the condition

$$y_i = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}}, \quad i = 1, \ldots, n$$

These equations are solvable for x if and only if $y \succ 0$ and $\mathbf{1}^T y = 1$. By substituting the expression for y_i into $y^T x - f(x)$ we obtain $f^*(y) = \sum_{i=1}^n y_i \log y_i$. This expression for f^* is still correct if some components of y are zero, as long as $y \succeq 0$ and $\mathbf{1}^T y = 1$ and we interpret $0 \log 0$ as 0.

In fact the domain of f^* is exactly given by $\mathbf{1}^T y = 1, y \succeq 0$. To show this, suppose that a component of y is negative, say, $y_k < 0$. Then we can show that $y^T x - f(x)$ is unbounded above by choosing $x_k = -t$, and $x_i = 0, i \neq k$, and letting t go to infinity. If $y \succeq 0$ but $\mathbf{1}^T y \neq 1$, we choose $x = t\mathbf{1}$, so that

$$y^T x - f(x) = t\mathbf{1}^T y - t - \log n$$

If $1^T y > 1$, this grows unboundedly as $t \to \infty$; if $1^T y < 1$, it grows unboundedly as $t \to -\infty$ In summary,

$$f^*(y) = \begin{cases} \sum_{i=1}^n y_i \log y_i & \text{if } y \succeq 0 \text{ and } \mathbf{1}^T y = 1 \\ \infty & \text{otherwise} \end{cases}$$

In other words, the conjugate of the log-sum-exp function is the negative entropy function, restricted to the probability simplex.

■ **Example 1.21 — Norm.** Let $\|\cdot\|$ be a norm on $\mathbf{R}^n$, with dual norm $\|\cdot\|_*$. We will show that the conjugate of $f(x) = \|x\|$ is

$$f^*(y) = \begin{cases} 0 & \|y\|_* \leq 1 \\ \infty & \text{otherwise} \end{cases}$$

i.e., the conjugate of a norm is the indicator function of the dual norm unit ball. If $\|y\|_* > 1$, then by definition of the dual norm, there is a $z \in \mathbf{R}^n$ with $\|z\| \leq 1$ and $y^T z > 1$. Taking $x = tz$ and letting $t \to \infty$, we have

$$y^T x - \|x\| = t\left(y^T z - \|z\|\right) \to \infty$$

which shows that $f^*(y) = \infty$. Conversely, if $\|y\|_* \leq 1$, then we have $y^T x \leq \|x\| \|y\|_*$ for all x, which implies for all $x, y^T x - \|x\| \leq 0$. Therefore $x = 0$ is the value that maximizes $y^T x - \|x\|$, with maximum value 0.

Basic properties

Theorem 1.3.12 — Fenchel's inequality.

$$f(x) + f^*(y) \geq x^T y$$

for all x, y. This is called **Fenchel's inequality** (or Young's inequality when f is differentiable).

■ **Example 1.22** For example with $f(x) = (1/2)x^T Qx$, where $Q \in \mathbf{S}^n_{++}$, we obtain the inequality

$$x^T y \leq (1/2)x^T Qx + (1/2)y^T Q^{-1} y$$

Theorem 1.3.13 — Conjugate of the conjugate. If f is convex, and f is closed (*i.e.*, epi f is a closed set), then $f^{**} = f$. For example, if $\operatorname{dom} f = \mathbf{R}^n$, then we have $f^{**} = f$, i.e., the conjugate of the conjugate of f is f again.

Definition 1.3.8 — Differentiable functions. The conjugate of a differentiable function f is also called the **Legendre transform** of f.

Suppose f is convex and differentiable, with $\operatorname{dom} f = \mathbf{R}^n$. Any maximizer x^* of $y^T x - f(x)$ satisfies $y = \nabla f(x^*)$, and conversely, if x^* satisfies $y = \nabla f(x^*)$, then x^* maximizes $y^T x - f(x)$. Therefore, if $y = \nabla f(x^*)$, we have

$$f^*(y) = x^{*T} \nabla f(x^*) - f(x^*)$$

This allows us to determine $f^*(y)$ for any y for which we can solve the gradient equation $y = \nabla f(z)$ for z

We can express this another way. Let $z \in \mathbf{R}^n$ be arbitrary and define $y = \nabla f(z)$. Then we have

$$f^*(y) = z^T \nabla f(z) - f(z)$$

Theorem 1.3.14 — Scaling and composition with affine transformation. For $a > 0$ and $b \in \mathbf{R}$, the conjugate of $g(x) = af(x) + b$ is $g^*(y) = af^*(y/a) - b$ Suppose $A \in \mathbf{R}^{n\times n}$ is nonsingular and $b \in \mathbf{R}^n$. Then the conjugate of $g(x) = f(Ax + b)$ is

$$g^*(y) = f^*\left(A^{-T} y\right) - b^T A^{-T} y$$

with $\operatorname{dom} g^* = A^T \operatorname{dom} f^*$

Theorem 1.3.15 — Sums of independent functions. If $f(u, v) = f_1(u) + f_2(v)$, where f_1

and f_2 are convex functions with conjugates f_1^* and f_2^*, respectively, then

$$f^*(w,z) = f_1^*(w) + f_2^*(z)$$

In other words, the conjugate of the sum of independent convex functions is the sum of the conjugates. ('Independent' means they are functions of different variables.)

1.3.7 Quasiconvex functions

Definition 1.3.9 — Quasiconvex.

1. A function $f : \mathbf{R}^n \to \mathbf{R}$ is called **quasiconvex** (or unimodal) if its domain and all its sublevel sets

$$S_\alpha = \{x \in \operatorname{dom} f \mid f(x) \le \alpha\}$$

for $\alpha \in \mathbf{R}$, are convex.
2. A function is **quasiconcave** if $-f$ is quasiconvex, *i.e.*, every superlevel set $\{x \mid f(x) \ge \alpha\}$ is convex.
3. A function that is both quasiconvex and quasiconcave is called **quasilinear**. If a function f is quasilinear, then its domain, and every level set $\{x \mid f(x) = \alpha\}$ is convex.

quasiconvexity is a considerable generalization of convexity.

Convex functions have convex sublevel sets, and so are quasiconvex. But simple examples, such as the one shown in figure 3.9, show that the converse is not true.

■ **Example 1.23 — Linear-fractional function.** The function

$$f(x) = \frac{a^T x + b}{c^T x + d}$$

with $\operatorname{dom} f = \left\{x \mid c^T x + d > 0\right\}$, is quasiconvex, and quasiconcave, *i.e.*, quasilinear. Its α -sublevel set is

$$\begin{aligned} S_\alpha &= \left\{x \mid c^T x + d > 0, \quad \left(a^T x + b\right) / \left(c^T x + d\right) \le \alpha\right\} \\ &= \left\{x \mid c^T x + d > 0, a^T x + b \le \alpha \left(c^T x + d\right)\right\} \end{aligned}$$

which is convex, since it is the intersection of an open halfspace and a closed halfspace. (The same method can be used to show its superlevel sets are convex.)

Theorem 1.3.16 — Jensen's inequality for quasiconvex functions. A function f is **quasiconvex** if and only if $\operatorname{dom} f$ is convex and for any $x, y \in \operatorname{dom} f$ and $0 \le \theta \le 1$

$$f(\theta x + (1-\theta)y) \le \max\{f(x), f(y)\} \tag{1.18}$$

i.e., the value of the function on a segment does not exceed the maximum of its values at the endpoints.

■ **Example 1.24 — Rank of positive semidefinite matrix.** The function rank X is quasiconcave on $\mathbf{S}_+^n$. This follows from the modified Jensen inequality (1.18),

$$\operatorname{rank}(X+Y) \geq \min\{\operatorname{rank} X, \operatorname{rank} Y\}$$

which holds for $X, Y \in \mathbf{S}_+^n$. (This can be considered an extension of the previous example, since $\operatorname{rank}(\operatorname{diag}(x)) = \operatorname{card}(x)$ for $x \succeq 0$.)

Theorem 1.3.17 f is quasiconvex if and only if its restriction to any line intersecting its domain is quasiconvex. In particular, quasiconvexity of a function can be verified by restricting it to an arbitrary line, and then checking quasiconvexity of the resulting function on $\mathbf{R}$.

Theorem 1.3.18 — Quasiconvex functions on R. A continuous function $f : \mathbf{R} \to \mathbf{R}$ is quasiconvex if and only if at least one of the following conditions holds:

1. f is nondecreasing
2. f is nonincreasing
3. there is a point $c \in \operatorname{dom} f$ such that for $t \leq c$ (and $t \in \operatorname{dom} f$), f is nonincreasing, and for $t \geq c$ (and $t \in \operatorname{dom} f$), f is nondecreasing.

The point c can be chosen as any point which is a global minimizer of f.

Theorem 1.3.19 — First-order conditions. Suppose $f : \mathbf{R}^n \to \mathbf{R}$ is differentiable. Then f is quasiconvex if and only if $\operatorname{dom} f$ is convex and for all $x, y \in \operatorname{dom} f$

$$f(y) \leq f(x) \Longrightarrow \nabla f(x)^T (y-x) \leq 0$$

Proposition 1.3.20 A simple geometric interpretation when $\nabla f(x) \neq 0 : \nabla f(x)$ defines a supporting hyperplane to the sublevel set $\{y \mid f(y) \leq f(x)\}$, at the point x.

Proposition 1.3.21 Difference between convex and quasiconvex: if f is convex and $\nabla f(x) = 0$, then x is a global minimizer of f. But this statement is false for quasiconvex functions: it is possible that $\nabla f(x) = 0$, but x is not a global minimizer of f.

Theorem 1.3.22 — Second-order conditions. Now suppose f is twice differentiable. If f is quasiconvex, then for all $x \in \operatorname{dom} f$, and all $y \in \mathbf{R}^n$, we have

$$y^T \nabla f(x) = 0 \Longrightarrow y^T \nabla^2 f(x) y \geq 0 \tag{1.19}$$

For a quasiconvex function on $\mathbf{R}$, this reduces to the simple condition

$$f'(x) = 0 \Longrightarrow f''(x) \geq 0$$

i.e., at any point with zero slope, the second derivative is nonnegative.

As in the case $n = 1$, we conclude that whenever $\nabla f(x) = 0$, we must have $\nabla^2 f(x) \succeq 0$. When $\nabla f(x) \neq 0$, the condition (1.19) means that $\nabla^2 f(x)$ is positive semidefinite on the $(n-1)$ -dimensional subspace $\nabla f(x)^\perp$. This implies that $\nabla^2 f(x)$ can have at most one negative eigenvalue.

As a (partial) converse, if f satisfies

$$y^T \nabla f(x) = 0 \Longrightarrow y^T \nabla^2 f(x) y > 0 \tag{1.20}$$

for all $x \in \operatorname{dom} f$ and all $y \in \mathbf{R}^n, y \neq 0$, then f is quasiconvex. This condition is the same as requiring $\nabla^2 f(x)$ to be positive definite for any point with $\nabla f(x) = 0$ and for all other points, requiring $\nabla^2 f(x)$ to be positive definite on the $(n-1)$ dimensional subspace $\nabla f(x)^\perp$.

Proof. By restricting the function to an arbitrary line, it suffices to consider the case in which $f : \mathbf{R} \to \mathbf{R}$

We first show that if $f : \mathbf{R} \to \mathbf{R}$ is quasiconvex on an interval (a, b), then it must satisfy (1.19), i.e., if $f'(c) = 0$ with $c \in (a, b)$, then we must have $f''(c) \geq 0$. If $f'(c) = 0$ with $c \in (a, b), f''(c) < 0$, then for small positive ϵ we have $f(c - \epsilon) < f(c)$ and $f(c + \epsilon) < f(c)$. It follows that the sublevel set $\{x \mid f(x) \leq f(c) - \epsilon\}$ is disconnected for small positive ϵ, and therefore not convex, which contradicts our assumption that f is quasiconvex.

Now we show that if the condition (1.20) holds, then f is quasiconvex. Assume that (1.20) holds, i.e., for each $c \in (a, b)$ with $f'(c) = 0$, we have $f''(c) > 0$. This means that whenever the function f' crosses the value 0, it is strictly increasing. Therefore it can cross the value 0 at most once. If f' does not cross the value 0 at all, then f is either nonincreasing or nondecreasing on (a, b), and therefore quasiconvex. Otherwise it must cross the value 0 exactly once, say at $c \in (a, b)$. Since $f''(c) > 0$, it follows that $f'(t) \leq 0$ for $a < t \leq c$, and $f'(t) \geq 0$ for $c \leq t < b$ This shows that f is quasiconvex. ■

Operations that preserve quasiconvexity

Theorem 1.3.23 — Nonnegative weighted maximum. A nonnegative weighted maximum of quasiconvex functions, i.e.,

$$f = \max\{w_1 f_1, \ldots, w_m f_m\}$$

with $w_i \geq 0$ and f_i quasiconvex, is quasiconvex. The property extends to the general pointwise supremum

$$f(x) = \sup_{y \in C} (w(y) g(x, y))$$

where $w(y) \geq 0$ and $g(x, y)$ is quasiconvex in x for each y. This fact can be easily verified: $f(x) \leq \alpha$ if and only if

$$w(y)g(x,y) \leq \alpha \text{ for all } y \in C$$

i.e., the α -sublevel set of f is the intersection of the α -sublevel sets of the functions $w(y)g(x,y)$ in the variable x

Theorem 1.3.24 — Composition. 1. If $g: \mathbf{R}^n \rightarrow \mathbf{R}$ is quasiconvex and $h: \mathbf{R} \rightarrow \mathbf{R}$ is nondecreasing, then $f = h \circ g$ is quasiconvex.

2. The composition of a quasiconvex function with an affine or linear-fractional transformation yields a quasiconvex function. If f is quasiconvex, then $g(x) = f(Ax+b)$ is quasiconvex, and $\tilde{g}(x) = f\left((Ax+b)/\left(c^T x+d\right)\right)$ is quasiconvex on the set

$$\left\{x \mid c^T x+d>0,(Ax+b)/\left(c^T x+d\right) \in \operatorname{dom} f\right\}$$

Theorem 1.3.25 — Minimization. If $f(x,y)$ is quasiconvex jointly in x and y and C is a convex set, then the function

$$g(x) = \inf_{y \in C} f(x,y)$$

is quasiconvex.

Proof. To show this, we need to show that $\{x \mid g(x) \leq \alpha\}$ is convex, where $\alpha \in \mathbf{R}$ is arbitrary. From the definition of $g, g(x) \leq \alpha$ if and only if for any $\epsilon>0$ there exists a $y \in C$ with $f(x,y) \leq \alpha+\epsilon$. Now let x_1 and x_2 be two points in the α -sublevel set of g. Then for any $\epsilon>0$, there exists $y_1, y_2 \in C$ with

$$f\left(x_1, y_1\right) \leq \alpha+\epsilon, \quad f\left(x_2, y_2\right) \leq \alpha+\epsilon$$

and since f is quasiconvex in x and y, we also have

$$f\left(\theta x_1+(1-\theta) x_2, \theta y_1+(1-\theta) y_2\right) \leq \alpha+\epsilon$$

for $0 \leq \theta \leq 1$. Hence $g\left(\theta x_1+(1-\theta) x_2\right) \leq \alpha$, which proves that $\{x \mid g(x) \leq \alpha\}$ is convex. ∎

Theorem 1.3.26 — Representation via family of convex functions. In the sequel, it will be convenient to represent the sublevel sets of a quasiconvex function f(which are convex) via inequalities of convex functions. We seek a family of convex functions $\phi_t: \mathbf{R}^n \rightarrow \mathbf{R}$, indexed by $t \in \mathbf{R}$, with

$$f(x) \leq t \Longleftrightarrow \phi_t(x) \leq 0$$

i.e., the t -sublevel set of the quasiconvex function f is the 0 -sublevel set of the convex function ϕ_t. Evidently ϕ_t must satisfy the property that for all $x \in \mathbf{R}^n$ $\phi_t(x) \leq 0 \Longrightarrow \phi_s(x) \leq 0$ for $s \geq t$. This is satisfied if for each $x, \phi_t(x)$ is a nonincreasing function of t, i.e., $\phi_s(x) \leq \phi_t(x)$ whenever $s \geq t$. To see that such a representation always exists, we can take

$$\phi_t(x) = \begin{cases} 0 & f(x) \leq t \\ \infty & \text{otherwise} \end{cases}$$

i.e., ϕ_t is the indicator function of the t -sublevel of f. Obviously this representation is not unique; for example if the sublevel sets of f are closed, we can take

$$\phi_t(x) = \operatorname{dist}(x, \{z \mid f(z) \leq t\})$$

■ **Example 1.25 — Convex over concave function.** Suppose p is a convex function, q is a concave function, with $p(x) \geq 0$ and $q(x) > 0$ on a convex set C. Then the function f defined by $f(x) = p(x)/q(x)$, on C, is quasiconvex. Here we have

$$f(x) \leq t \Longleftrightarrow p(x) - tq(x) \leq 0$$

so we can take $\phi_t(x) = p(x) - tq(x)$ for $t \geq 0$. For each t, ϕ_t is convex and for each $x, \phi_t(x)$ is decreasing in t.

1.3.8 Log-concave and log-convex functions

Definition 1.3.10
1. A function $f: \mathbf{R}^n \to \mathbf{R}$ is **logarithmically concave or log-concave** if $f(x) > 0$ for all $x \in \operatorname{dom} f$ and $\log f$ is concave.
2. It is said to be logarithmically convex or log-convex if $\log f$ is convex. Thus f is log-convex if and only if $1/f$ is logconcave.
3. It is convenient to allow f to take on the value zero, in which case we take $\log f(x) = -\infty$. In this case we say f is log-concave if the extended-value function $\log f$ is concave.

Theorem 1.3.27 — Express log-concavity directly, without logarithms. a function $f: \mathbf{R}^n \to \mathbf{R}$, with convex domain and $f(x) > 0$ for all $x \in \operatorname{dom} f$, is log-concave if and only if for all $x, y \in \operatorname{dom} f$ and $0 \leq \theta \leq 1$, we have

$$f(\theta x + (1-\theta)y) \geq f(x)^\theta f(y)^{1-\theta}$$

In particular, the value of a log-concave function at the average of two points is at least the geometric mean of the values at the two points.

Proposition 1.3.28
1. From the composition rules we know that e^h is convex if h is convex, so a logconvex function is convex.

2. Similarly, a nonnegative concave function is log-concave.
3. It is also clear that a log-convex function is quasiconvex and a log-concave function is quasiconcave, since the logarithm is monotone increasing.

■ **Example 1.26 — Log-concave density functions.** Many common probability density functions are log-concave. Two examples are the multivariate normal distribution,

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} e^{-\frac{1}{2}(x-\bar{x})^T \Sigma^{-1}(x-\bar{x})}$$

(where $\bar{x} \in \mathbf{R}^n$ and $\Sigma \in \mathbf{S}^n_{++}$), and the exponential distribution on $\mathbf{R}^n_+$,

$$f(x) = \left(\prod_{i=1}^n \lambda_i\right) e^{-\lambda^T x}$$

(where $\lambda \succ 0$). Another example is the uniform distribution over a convex set C,

$$f(x) = \begin{cases} 1/\alpha & x \in C \\ 0 & x \notin C \end{cases}$$

where $\alpha = \mathrm{vol}(C)$ is the volume (Lebesgue measure) of C. In this case $\log f$ takes on the value $-\infty$ outside C, and $-\log \alpha$ on C, hence is concave.

As a more exotic example consider the Wishart distribution, defined as follows. Let $x_1, \ldots, x_p \in \mathbf{R}^n$ be independent Gaussian random vectors with zero mean and covariance $\Sigma \in \mathbf{S}^n$, with $p > n$. The random matrix $X = \sum_{i=1}^p x_i x_i^T$ has the Wishart density

$$f(X) = a(\det X)^{(p-n-1)/2} e^{-\frac{1}{2}\mathrm{tr}\left(\Sigma^{-1}X\right)}$$

with $\mathrm{dom}\, f = \mathbf{S}^n_{++}$, and a is a positive constant. The Wishart density is log-concave, since

$$\log f(X) = \log a + \frac{p-n-1}{2} \log\det X - \frac{1}{2}\mathrm{tr}\left(\Sigma^{-1}X\right)$$

which is a concave function of X.

Properties

Proposition 1.3.29 — Twice differentiable log-convex/concave functions. Suppose f is twice differentiable, with $\mathrm{dom}\, f$ convex, so

$$\nabla^2 \log f(x) = \frac{1}{f(x)} \nabla^2 f(x) - \frac{1}{f(x)^2} \nabla f(x) \nabla f(x)^T$$

We conclude that f is log-convex if and only if for all $x \in \mathrm{dom}\, f$,

$$f(x)\nabla^2 f(x) \succeq \nabla f(x) \nabla f(x)^T$$

and log-concave if and only if for all $x \in \mathrm{dom}\ f$,

$$f(x)\nabla^2 f(x) \preceq \nabla f(x) \nabla f(x)^T$$

Proposition 1.3.30 — **Log-convexity and log-concavity are closed under multiplication and positive scaling.** For example, if f and g are log-concave, then so is the pointwise product $h(x) = f(x)g(x)$, since $\log h(x) = \log f(x) + \log g(x)$, and $\log f(x)$ and $\log g(x)$ are concave functions of x

Proposition 1.3.31 — **The sum of log-concave functions is not, in general, log-concave. Log-convexity, however, is preserved under sums.** Let f and g be logconvex functions, i.e., $F = \log f$ and $G = \log g$ are convex. From the composition rules for convex functions, it follows that

$$\log(\exp F + \exp G) = \log(f + g)$$

is convex. Therefore the sum of two log-convex functions is log-convex. More generally, if $f(x,y)$ is log-convex in x for each $y \in C$ then

$$g(x) = \int_C f(x,y)dy$$

is log-convex.

■ **Example 1.27 — Laplace transform of a nonnegative function and the moment and cumulant generating functions.** Suppose $p : \mathbf{R}^n \to \mathbf{R}$ satisfies $p(x) \geq 0$ for all x. The Laplace transform of p

$$P(z) = \int p(x)e^{-z^T x}dx$$

is log-convex on $\mathbf{R}^n$. (Here dom P is, naturally, $\{z \mid P(z) < \infty\}$.) Now suppose p is a density, *i.e.*, satisfies $\int p(x)dx = 1$. The function $M(z) = P(-z)$ is called the **moment generating function** of the density. It gets its name from the fact that the moments of the density can be found from the derivatives of the moment generating function, evaluated at $z = 0$, e.g.

$$\nabla M(0) = \mathbf{E}v, \quad \nabla^2 M(0) = \mathbf{E}vv^T$$

where v is a random variable with density p. The function $\log M(z)$, which is convex, is called the **cumulant generating function** for p, since its derivatives give the cumulants of the density. For example, the first and second derivatives of the cumulant generating function, evaluated at zero, are the mean and covariance of the associated random variable:

$$\nabla \log M(0) = \mathbf{E}v, \quad \nabla^2 \log M(0) = \mathbf{E}(v - \mathbf{E}v)(v - \mathbf{E}v)^T$$

Integration of log-concave functions

Proposition 1.3.32 — **log-concavity is preserved by integration.** If $f : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}$ is log-concave, then

$$g(x) = \int f(x,y)dy$$

is a log-concave function of x (on $\mathbf{R}^n$). (The integration here is over $\mathbf{R}^m$.)

This result implies that marginal distributions (integral one variable) of logconcave probability densities are log-concave. It also implies that log-concavity is closed under convolution, i.e., if f and g are log-concave on $\mathbf{R}^n$, then so is the convolution

$$(f*g)(x) = \int f(x-y)g(y)dy$$

(To see this, note that $g(y)$ and $f(x-y)$ are log-concave in (x,y), hence the product $f(x-y)g(y)$ is; then the integration result applies.)

■ **Example 1.28** Suppose $C \subseteq \mathbf{R}^n$ is a convex set and w is a random vector in $\mathbf{R}^n$ with logconcave probability density p. Then the function

$$f(x) = \mathrm{prob}(x+w \in C)$$

is log-concave in x. To see this, express f as

$$f(x) = \int g(x+w)p(w)dw$$

where g is defined as

$$g(u) = \begin{cases} 1 & u \in C \\ 0 & u \notin C \end{cases}$$

(which is log-concave) and apply the integration result.

■ **Example 1.29** The cumulative distribution function of a probability density function $f : \mathbf{R}^n \to \mathbf{R}$ is defined as

$$F(x) = \mathrm{prob}(w \preceq x) = \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} f(z)dz_1 \cdots dz_n$$

where w is a random variable with density f. If f is log-concave, then F is logconcave. We have already encountered a special case: the cumulative distribution function of a Gaussian random variable,

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}dt$$

is log-concave.

■ **Example 1.30 — Yield function.** The yield of the manufacturing process, as a function of the nominal parameter values, is given by

$$Y(x) = \mathrm{prob}(x+w \in S)$$

For example, the 95% yield region

$$\{x \mid Y(x) \geq 0.95\} = \{x \mid \log Y(x) \geq \log 0.95\}$$

is convex, since it is a superlevel set of the concave function $\log Y$.

Definition 1.3.11 — Monotonicity with respect to a generalized inequality. Suppose $K \subseteq \mathbf{R}^n$ is a proper cone with associated generalized inequality $\preceq_K$. A function $f : \mathbf{R}^n \to \mathbf{R}$ is called K **-nondecreasing** if

$$x \preceq_K y \Longrightarrow f(x) \leq f(y)$$

and K **-increasing** if

$$x \preceq_K y, x \neq y \Longrightarrow f(x) < f(y)$$

We define K -nonincreasing and K -decreasing functions in a similar way.

■ **Example 1.31 — Monotone vector functions.** A function $f : \mathbf{R}^n \to \mathbf{R}$ is nondecreasing with respect to $\mathbf{R}^n_+$ if and only if

$$x_1 \leq y_1, \ldots, x_n \leq y_n \Longrightarrow f(x) \leq f(y)$$

for all x, y. This is the same as saying that f, when restricted to any component x_i (i.e., x_i is considered the variable while x_j for $j \neq i$ are fixed), is nondecreasing.

■ **Example 1.32 — Matrix monotone functions.** A function $f : \mathbf{S}^n \to \mathbf{R}$ is called matrix monotone (increasing, decreasing) if it is monotone with respect to the positive semidefinite cone. Some examples of matrix monotone functions of the variable $X \in \mathbf{S}^n$:

1. $\mathbf{tr}(WX)$, where $W \in \mathbf{S}^n$, is matrix nondecreasing if $W \succeq 0$, and matrix increasing if $W \succ 0$ (it is matrix nonincreasing if $W \preceq 0$, and matrix decreasing if $W \prec 0$)
2. $\operatorname{tr}\left(X^{-1}\right)$ is matrix decreasing on $\mathbf{S}^n_{++}$
3. $\det X$ is matrix increasing on $\mathbf{S}^n_{++}$, and matrix nondecreasing on $\mathbf{S}^n_+$

Proposition 1.3.33 — **Gradient conditions for monotonicity.** A differentiable function f, with convex domain, is K **-nondecreasing** if and only if

$$\nabla f(x) \succeq_{K^*} 0 \tag{1.21}$$

for all $x \in \operatorname{dom} f$. Note the difference with the simple scalar case: the gradient must be nonnegative in the *dual* inequality. For the strict case, we have the following: If

$$\nabla f(x) \succ_{K^*} 0$$

for all $x \in \operatorname{dom} f$, then f is K -increasing. As in the scalar case, the converse is not true.

Proof. Let us prove these first-order conditions for monotonicity. First, assume that f satisfies (1.21) for all x, but is not K -nondecreasing, *i.e.*, there exist x, y with $x \preceq_K y$ and $f(y) < f(x)$. By differentiability of f there exists a $t \in [0,1]$ with

$$\frac{d}{dt} f(x + t(y - x)) = \nabla f(x + t(y - x))^T (y - x) < 0$$

since $y - x \in K$ this means

$$\nabla f(x + t(y - x)) \notin K^*$$

which contradicts our assumption that (1.21) is satisfied everywhere.

It is also straightforward to see that it is necessary that (1.21) hold everywhere. Assume (1.21) does not hold for $x = z$. By the definition of dual cone this means there exists a $v \in K$ with

$$\nabla f(z)^T v < 0$$

Now consider $h(t) = f(z + tv)$ as a function of t. We have $h'(0) = \nabla f(z)^T v < 0$ and therefore there exists $t > 0$ with $h(t) = f(z + tv) < h(0) = f(z)$, which means f is not K -nondecreasing. ■

Definition 1.3.12 — Convexity with respect to a generalized inequality. Suppose $K \subseteq \mathbf{R}^m$ is a proper cone with associated generalized inequality $\preceq_K$. We say $f : \mathbf{R}^n \to \mathbf{R}^m$ is K **-convex** if for all x, y, and $0 \le \theta \le 1$,

$$f(\theta x + (1 - \theta) y) \preceq_K \theta f(x) + (1 - \theta) f(y)$$

The function is **strictly** K **-convex** if

$$f(\theta x + (1 - \theta) y) \prec_K \theta f(x) + (1 - \theta) f(y)$$

for all $x \neq y$ and $0 < \theta < 1$. These definitions reduce to ordinary convexity and strict convexity when $m = 1$ (and $K = \mathbf{R}_+$)

■ **Example 1.33 — Matrix convexity.** Suppose f is a symmetric matrix valued function, i.e., $f : \mathbf{R}^n \to \mathbf{S}^m$. The function f is convex with respect to matrix inequality if

$$f(\theta x + (1 - \theta) y) \preceq \theta f(x) + (1 - \theta) f(y)$$

for any x and y, and for $\theta \in [0, 1]$. This is sometimes called **matrix convexity**. An equivalent definition is that the scalar function $z^T f(x) z$ is convex for all vectors z. (This is often a good way to prove matrix convexity). A matrix function is strictly matrix convex if

$$f(\theta x + (1 - \theta) y) \prec \theta f(x) + (1 - \theta) f(y)$$

when $x \neq y$ and $0 < \theta < 1$, or, equivalently, if $z^T f z$ is strictly convex for every $z \neq 0$ Some examples:

1. The function $f(X) = XX^T$ where $X \in \mathbf{R}^{n \times m}$ is matrix convex, since for fixed z the function $z^T XX^T z = \|X^T z\|_2^2$ is a convex quadratic function of (the components of) X. For the same reason, $f(X) = X^2$ is matrix convex on $\mathbf{S}^n$.

2. The function X^p is matrix convex on $\mathbf{S}^n_{++}$ for $1 \leq p \leq 2$ or $-1 \leq p \leq 0$, and matrix concave for $0 \leq p \leq 1$
3. The function $f(X) = e^X$ is not matrix convex on $\mathbf{S}^n$, for $n \geq 2$

Proposition 1.3.34 A function is K **-convex** if and only if its restriction to any line in its domain is K -convex.

Definition 1.3.13 — Dual characterization of K -convexity.
1. A function f is K -convex if and only if for every $w \succeq_{K*} 0$, the (real-valued) function $w^T f$ is convex (in the ordinary sense);
2. f is strictly K -convex if and only if for every nonzero $w \succeq_{K^*} 0$ the function $w^T f$ is strictly convex. (These follow directly from the definitions and properties of dual inequality.)

Proposition 1.3.35 — Differentiable K -convex functions. A differentiable function f is K -convex if and only if its domain is convex, and for all $x, y \in \operatorname{dom} f$

$$f(y) \succeq_K f(x) + Df(x)(y - x)$$

(Here $Df(x) \in \mathbf{R}^{m \times n}$ is the derivative or Jacobian matrix of f at x). The function f is strictly K -convex if and only if for all $x, y \in \operatorname{dom} f$ with $x \neq y$,

$$f(y) \succ_K f(x) + Df(x)(y - x)$$

Proposition 1.3.36 — Composition theorem.
1. if $g : \mathbf{R}^n \to \mathbf{R}^p$ is K -convex, $h : \mathbf{R}^p \to \mathbf{R}$ is convex, and $\tilde{h}$ (the extended-value extension of h) is K -nondecreasing, then $h \circ g$ is convex.
2. This generalizes the fact that a nondecreasing convex function of a convex function is convex. The condition that $\tilde{h}$ be K -nondecreasing implies that $\operatorname{dom} h - K = \operatorname{dom} h$

■ **Example 1.34** The quadratic matrix function $g : \mathbf{R}^{m \times n} \to \mathbf{S}^n$ defined by

$$g(X) = X^T A X + B^T X + X^T B + C$$

where $A \in \mathbf{S}^m, B \in \mathbf{R}^{m \times n}$, and $C \in \mathbf{S}^n$, is convex when $A \succeq 0$. The function $h : \mathbf{S}^n \to \mathbf{R}$ defined by $h(Y) = -\log\det(-Y)$ is convex and increasing on $\operatorname{dom} h = -\mathbf{S}^n_{++}$. By the composition theorem, we conclude that

$$f(X) = -\log\det\left(-\left(X^T A X + B^T X + X^T B + C\right)\right)$$

is convex on

$$\operatorname{dom} f = \left\{X \in \mathbf{R}^{m \times n} \mid X^T A X + B^T X + X^T B + C \prec 0\right\}$$

This generalizes the fact that

$$-\log\left(-\left(ax^2 + bx + c\right)\right)$$

is convex on

$$\left\{x \in \mathbf{R} \mid a x^2+b x+c<0\right\}$$

provided $a \geq 0$

1.4 Convex optimization problems

1.4.1 Optimization problems

We use the notation

$$\begin{array}{ll}
\operatorname{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \leq 0, \quad i=1, \ldots, m \\
& h_i(x)=0, \quad i=1, \ldots, p
\end{array} \tag{1.22}$$

to describe the problem of finding an x that minimizes $f_0(x)$ among all x that satisfy the conditions $f_i(x) \leq 0, i=1, \ldots, m$, and $h_i(x)=0, i=1, \ldots, p$. We call $x \in \mathbf{R}^n$ the optimization variable and the function $f_0: \mathbf{R}^n \rightarrow \mathbf{R}$ the objective function or cost function. The inequalities $f_i(x) \leq 0$ are called inequality constraints, and the corresponding functions $f_i: \mathbf{R}^n \rightarrow \mathbf{R}$ are called the inequality constraint functions. The equations $h_i(x)=0$ are called the equality constraints, and the functions $h_i: \mathbf{R}^n \rightarrow \mathbf{R}$ are the equality constraint functions. If there are no constraints (*i.e.*, $m=p=0$) we say the problem (1.22) is unconstrained.

The set of points for which the objective and all constraint functions are defined,

$$\mathcal{D}=\bigcap_{i=0}^m \operatorname{dom} f_i \cap \bigcap_{i=1}^p \operatorname{dom} h_i$$

is called the domain of the optimization problem (4.1). A point $x \in \mathcal{D}$ is feasible if it satisfies the constraints $f_i(x) \leq 0, i=1, \ldots, m$, and $h_i(x)=0, i=1, \ldots, p$ The problem (4.1) is said to be feasible if there exists at least one feasible point, and infeasible otherwise. The set of all feasible points is called the feasible set or the constraint set. The optimal value $p^{\star}$ of the problem (4.1) is defined as

$$p^{\star}=\inf \left\{f_0(x) \mid f_i(x) \leq 0, i=1, \ldots, m, h_i(x)=0, i=1, \ldots, p\right\}$$

We allow $p^{\star}$ to take on the extended values $\pm \infty$. If the problem is infeasible, we have $p^{\star}=\infty$ (following the standard convention that the infimum of the empty set is ∞). If there are feasible points x_k with $f_0\left(x_k\right) \rightarrow-\infty$ as $k \rightarrow \infty$, then $p^{\star}=-\infty$ and we say the problem (1.22) is unbounded below...

Optimal and locally optimal points

We say $x^\star$ is an optimal point, or solves the problem (4.1), if $x^\star$ is feasible and $f_0(x^\star) = p^\star$. The set of all optimal points is the optimal set, denoted

$$X_{\text{opt}} = \{x \mid f_i(x) \leq 0, i = 1, \ldots, m, h_i(x) = 0, i = 1, \ldots, p, f_0(x) = p^\star\}$$

If there exists an optimal point for the problem (1.22), we say the optimal value is attained or achieved, and the problem is solvable. If X_{opt} is empty, we say the optimal value is not attained or not achieved. (This always occurs when the problem is unbounded below.) A feasible point x with $f_0(x) \leq p^\star + \epsilon$ (where $\epsilon > 0$) is called ϵ -suboptimal, and the set of all ϵ -suboptimal points is called the $\epsilon-$suboptimal set for the problem (1.22).

We say a feasible point x is locally optimal if there is an $R > 0$ such that

$$\begin{aligned} f_0(x) = \inf\{&f_0(z) \mid f_i(z) \leq 0, i = 1, \ldots, m \\ &h_i(z) = 0, i = 1, \ldots, p, \|z - x\|_2 \leq R\} \end{aligned}$$

or, in other words, x solves the optimization problem

$$\begin{array}{ll} \text{minimize} & f_0(z) \\ \text{subject to} & f_i(z) \leq 0, \quad i = 1, \ldots, m \\ & h_i(z) = 0, \quad i = 1, \ldots, p \\ & \|z - x\|_2 \leq R \end{array}$$

with variable z. Roughly speaking, this means x minimizes f_0 over nearby points in the feasible set. The term 'globally optimal' is sometimes used for 'optimal' to distinguish between 'locally optimal' and 'optimal'. Throughout this book, however, optimal will mean globally optimal.

If x is feasible and $f_i(x) = 0$, we say the i th inequality constraint $f_i(x) \leq 0$ is active at x. If $f_i(x) < 0$, we say the constraint $f_i(x) \leq 0$ is inactive. (The equality constraints are active at all feasible points.) We say that a constraint is redundant if deleting it does not change the feasible set.

1.4.2 Feasibility problems

If the objective function is identically zero, the optimal value is either zero (if the feasible set is nonempty) or ∞ (if the feasible set is empty). We call this the feasibility problem, and will sometimes write it as

$$\begin{array}{ll} \text{find} & x \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p \end{array}$$

The feasibility problem is thus to determine whether the constraints are consistent, and if so, find a point that satisfies them.

Expressing problems in standard form We refer to (4.1) as an optimization problem in standard form. In the standard form problem we adopt the convention that the righthand side of the inequality and equality constraints are zero. This can always be arranged by subtracting any nonzero righthand side: we represent the equality constraint $g_i(x) = \tilde{g}_i(x)$, for example, as $h_i(x) = 0$, where $h_i(x) = g_i(x) - \tilde{g}_i(x)$. In a similar way we express inequalities of the form $f_i(x) \geq 0$ as $-f_i(x) \leq 0$.

■ Example 1.35 — Box constraints. Consider the optimization problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & l_i \leq x_i \leq u_i, \quad i = 1, \ldots, n, \end{array}$$

where $x \in \mathbf{R}^n$ is the variable. The constraints are called variable bounds (since they give lower and upper bounds for each x_i) or box constraints (since the feasible set is a box) We can express this problem in standard form as

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & l_i - x_i \leq 0, \quad i = 1, \ldots, n \\ & x_i - u_i \leq 0, \quad i = 1, \ldots, n \end{array}$$

There are $2n$ inequality constraint functions:

$$f_i(x) = l_i - x_i, \quad i = 1, \ldots, n$$

and

$$f_i(x) = x_{i-n} - u_{i-n}, \quad i = n+1, \ldots, 2n$$

Maximization problems We concentrate on the minimization problem by convention. We can solve the maximization problem

$$\begin{array}{ll} \text{maximize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p \end{array} \tag{1.23}$$

by minimizing the function $-f_0$ subject to the constraints. By this correspondence we can define all the terms above for the maximization problem (1.23). For example the optimal value of (1.23) is defined as

$$p^\star = \sup \{ f_0(x) \mid f_i(x) \leq 0, i = 1, \ldots, m, h_i(x) = 0, i = 1, \ldots, p \}$$

and a feasible point x is ϵ -suboptimal if $f_0(x) \geq p^\star - \epsilon$. When the maximization problem is considered, the objective is sometimes called the utility or satisfaction level instead of the cost.

Equivalent problems

In this book we will use the notion of equivalence of optimization problems in an informal way. We call two problems equivalent if from a solution of one, a solution of the other is readily found, and vice versa. (It is possible, but complicated, to give a formal definition of equivalence.) As a simple example, consider the problem

$$\begin{array}{ll} \text{minimize} & \tilde{f}(x) = \alpha_0 f_0(x) \\ \text{subject to} & \tilde{f}_i(x) = \alpha_i f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & \tilde{h}_i(x) = \beta_i h_i(x) = 0, \quad i = 1, \ldots, p \end{array} \tag{1.24}$$

where $\alpha_i > 0, i = 0, \ldots, m$, and $\beta_i \neq 0, i = 1, \ldots, p$. This problem is obtained from the standard form problem (1.22) by scaling the objective and inequality constraint functions by positive constants, and scaling the equality constraint functions by nonzero constants. As a result, the feasible sets of the problem (1.24) and the original problem (1.22) are identical. A point x is optimal for the original problem (1.22) if and only if it is optimal for the scaled problem (1.24), so we say the two problems are equivalent. The two problems (1.22) and (1.24) are not, however, the same (unless α_i and β_i are all equal to one), since the objective and constraint functions differ. We now describe some general transformations that yield equivalent problems.

Change of variables

Suppose $\phi : \mathbf{R}^n \to \mathbf{R}^n$ is one-to-one, with image covering the problem domain $\mathcal{D}$ i.e., $\phi(\mathbf{dom}\,\phi) \supseteq \mathcal{D}$. We define functions $\tilde{f}_i$ and $\tilde{h}_i$ as

$$\tilde{f}_i(z) = f_i(\phi(z)), \quad i = 0, \ldots, m, \quad \tilde{h}_i(z) = h_i(\phi(z)), \quad i = 1, \ldots, p$$

Now consider the problem

$$\begin{array}{ll} \text{minimize} & \tilde{f}_0(z) \\ \text{subject to} & \tilde{f}_i(z) \leq 0, \quad i = 1, \ldots, m \\ & \tilde{h}_i(z) = 0, \quad i = 1, \ldots, p \end{array} \tag{1.25}$$

with variable z. We say that the standard form problem (1.22) and the problem (1.25) are related by the change of variable or substitution of variable $x = \phi(z)$.

The two problems are clearly equivalent: if x solves the problem (1.22), then $z = \phi^{-1}(x)$ solves the problem (1.25); if z solves the problem (1.25), then $x = \phi(z)$ solves the problem (1.22) .

Transformation of objective and constraint functions

Suppose that $\psi_0 : \mathbf{R} \to \mathbf{R}$ is monotone increasing, $\psi_1, \ldots, \psi_m : \mathbf{R} \to \mathbf{R}$ satisfy $\psi_i(u) \leq 0$ if and only if $u \leq 0$, and $\psi_{m+1}, \ldots, \psi_{m+p} : \mathbf{R} \to \mathbf{R}$ satisfy $\psi_i(u) = 0$ if and only if $u = 0$. We define functions $\tilde{f}_i$ and $\tilde{h}_i$ as the compositions

$$\tilde{f}_i(x) = \psi_i(f_i(x)), \quad i = 0, \ldots, m, \quad \tilde{h}_i(x) = \psi_{m+i}(h_i(x)), \quad i = 1, \ldots, p$$

Evidently the associated problem

$$\begin{array}{ll} \text{minimize} & \tilde{f}_0(x) \\ \text{subject to} & \tilde{f}_i(x) \leq 0, \quad i = 1, \dots, m \\ & \tilde{h}_i(x) = 0, \quad i = 1, \dots, p \end{array}$$

and the standard form problem (4.1) are equivalent; indeed, the feasible sets are identical, and the optimal points are identical. (The example (4.3) above, in which the objective and constraint functions are scaled by appropriate constants, is the special case when all ψ_i are linear.)

■ **Example 1.36** Least-norm and least-norm-squared problems. As a simple example consider the unconstrained Euclidean norm minimization problem

$$\text{minimize} \quad \|Ax - b\|_2 \tag{1.26}$$

with variable $x \in \mathbf{R}^n$. since the norm is always nonnegative, we can just as well solve the problem

$$\text{minimize} \quad \|Ax - b\|_2^2 = (Ax - b)^T (Ax - b) \,, \tag{1.27}$$

in which we minimize the square of the Euclidean norm. The problems (1.26) and (1.27) are clearly equivalent; the optimal points are the same. The two problems are not the same, however. For example, the objective in (1.27) is not differentiable at any x with $Ax - b = 0$, whereas the objective in (1.27) is differentiable for all x (in fact, quadratic).

Slack variables

One simple transformation is based on the observation that $f_i(x) \leq 0$ if and only if there is an $s_i \geq 0$ that satisfies $f_i(x) + s_i = 0$. Using this transformation we obtain the problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & s_i \geq 0, \quad i = 1, \dots, m \\ & f_i(x) + s_i = 0, \quad i = 1, \dots, m \\ & h_i(x) = 0, \quad i = 1, \dots, p \end{array} \tag{1.28}$$

where the variables are $x \in \mathbf{R}^n$ and $s \in \mathbf{R}^m$. This problem has $n + m$ variables, m inequality constraints (the nonnegativity constraints on s_i), and $m + p$ equality constraints. The new variable s_i is called the slack variable associated with the original inequality constraint $f_i(x) \leq 0$. Introducing slack variables replaces each inequality constraint with an equality constraint, and a nonnegativity constraint. The problem (1.28) is equivalent to the original standard form problem (1.22) . Indeed, if (x, s) is feasible for the problem (1.28), then x is feasible for the original problem, since $s_i = -f_i(x) \geq 0$. Conversely, if x is feasible for the original problem, then (x, s) is feasible for the problem (1.28), where we take $s_i = -f_i(x)$. Similarly, x is optimal for the original problem (4.1) if and only if (x, s) is optimal for the problem (1.28), where $s_i = -f_i(x)$

Eliminating equality constraints

If we can explicitly parametrize all solutions of the equality constraints

$$h_i(x) = 0, \quad i = 1, \ldots, p \tag{1.29}$$

using some parameter $z \in \mathbf{R}^k$, then we can eliminate the equality constraints from the problem, as follows. Suppose the function $\phi : \mathbf{R}^k \to \mathbf{R}^n$ is such that x satisfies (1.29) if and only if there is some $z \in \mathbf{R}^k$ such that $x = \phi(z)$. The optimization problem

$$\begin{array}{ll} \text{minimize} & \tilde{f}_0(z) = f_0(\phi(z)) \\ \text{subject to} & \tilde{f}_i(z) = f_i(\phi(z)) \leq 0, \quad i = 1, \ldots, m \end{array}$$

is then equivalent to the original problem (1.22). This transformed problem has variable $z \in \mathbf{R}^k, m$ inequality constraints, and no equality constraints. If z is optimal for the transformed problem, then $x = \phi(z)$ is optimal for the original conversely, if x is optimal for the original problem, then (since x is feasible) there is at least one z such that $x = \phi(z)$. Any such z is optimal for the transformed problem.

Eliminating linear equality constraints

The process of eliminating variables can be described more explicitly, and easily carried out numerically, when the equality constraints are all linear, *i.e.*, have the form $Ax = b$. If $Ax = b$ is inconsistent, i.e.,$b \notin \mathcal{R}(A)$, then the original problem is infeasible. Assuming this is not the case, let x_0 denote any solution of the equality constraints. Let $F \in \mathbf{R}^{n \times k}$ be any matrix with $\mathcal{R}(F) = \mathcal{N}(A)$, so the general solution of the linear equations $Ax = b$ is given by $Fz + x_0$, where $z \in \mathbf{R}^k$. (We can choose F to be full rank, in which case we have $k = n - \operatorname{rank} A$.) Substituting $x = Fz + x_0$ into the original problem yields the problem

$$\begin{array}{ll} \text{minimize} & f_0\left(Fz + x_0\right) \\ \text{subject to} & f_i\left(Fz + x_0\right) \leq 0, \quad i = 1, \ldots, m \end{array}$$

with variable z, which is equivalent to the original problem, has no equality constraints, and rank A fewer variables.

Introducing equality constraints

We can also introduce equality constraints and new variables into a problem. Instead of describing the general case, which is complicated and not very illuminating, we give a typical example that will be useful later. Consider the problem

$$\begin{array}{ll} \text{minimize} & f_0\left(A_0 x + b_0\right) \\ \text{subject to} & f_i\left(A_i x + b_i\right) \leq 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p \end{array}$$

where $x \in \mathbf{R}^n, A_i \in \mathbf{R}^{k_i \times n}$, and $f_i : \mathbf{R}^{k_i} \to \mathbf{R}$. In this problem the objective and constraint functions are given as compositions of the functions f_i with affine transformations defined by $A_i x + b_i$

We introduce new variables $y_i \in \mathbf{R}^{k_i}$, as well as new equality constraints $y_i = A_i x + b_i$, for $i = 0, \ldots, m$, and form the equivalent problem

$$\begin{array}{ll} \text{minimize} & f_0(y_0) \\ \text{subject to} & f_i(y_i) \leq 0, \quad i = 1, \ldots, m \\ & y_i = A_i x + b_i, \quad i = 0, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p \end{array}$$

This problem has $k_0 + \cdots + k_m$ new variables,

$$y_0 \in \mathbf{R}^{k_0}, \quad \ldots, \quad y_m \in \mathbf{R}^{k_m}$$

and $k_0 + \cdots + k_m$ new equality constraints,

$$y_0 = A_0 x + b_0, \quad \ldots, \quad y_m = A_m x + b_m$$

The objective and inequality constraints in this problem are independent, i.e., involve different optimization variables.

Optimizing over some variables We always have

$$\inf_{x,y} f(x, y) = \inf_x \tilde{f}(x)$$

where $\tilde{f}(x) = \inf_y f(x, y)$. In other words, we can always minimize a function by first minimizing over some of the variables, and then minimizing over the remaining ones. This simple and general principle can be used to transform problems into equivalent forms. The general case is cumbersome to describe and not illuminating, so we describe instead an example.

Suppose the variable $x \in \mathbf{R}^n$ is partitioned as $x = (x_1, x_2)$, with $x_1 \in \mathbf{R}^{n_1}$, $x_2 \in \mathbf{R}^{n_2}$, and $n_1 + n_2 = n$. We consider the problem

$$\begin{array}{ll} \text{minimize} & f_0(x_1, x_2) \\ \text{subject to} & f_i(x_1) \leq 0, \quad i = 1, \ldots, m_1 \\ & \tilde{f}_i(x_2) \leq 0, \quad i = 1, \ldots, m_2 \end{array} \tag{1.30}$$

in which the constraints are independent, in the sense that each constraint function depends on x_1 or x_2. We first minimize over x_2. Define the function $\tilde{f}_0$ of x_1 by

$$\tilde{f}_0(x_1) = \inf\{f_0(x_1, z) \mid \tilde{f}_i(z) \leq 0, i = 1, \ldots, m_2\}$$

The problem (1.30) is then equivalent to

$$\begin{array}{ll} \text{minimize} & \tilde{f}_0(x_1) \\ \text{subject to} & f_i(x_1) \leq 0, \quad i = 1, \ldots, m_1 \end{array} \tag{1.31}$$

■ **Example 1.37** Minimizing a quadratic function with constraints on some variables. Consider a problem with strictly convex quadratic objective, with some of the variables unconstrained:

$$\begin{array}{ll} \text{minimize} & x_1^T P_{11} x_1 + 2x_1^T P_{12} x_2 + x_2^T P_{22} x_2 \\ \text{subject to} & f_i(x_1) \leq 0, \quad i = 1, \ldots, m \end{array}$$

where P_{11} and P_{22} are symmetric. Here we can analytically minimize over x_2 :

$$\inf_{x_2} \left(x_1^T P_{11} x_1 + 2x_1^T P_{12} x_2 + x_2^T P_{22} x_2 \right) = x_1^T \left(P_{11} - P_{12} P_{22}^{-1} P_{12}^T \right) x_1$$

(see §A.5.5). Therefore the original problem is equivalent to

$$\begin{array}{ll} \text{minimize} & x_1^T \left(P_{11} - P_{12} P_{22}^{-1} P_{12}^T \right) x_1 \\ \text{subject to} & f_i(x_1) \leq 0, \quad i = 1, \ldots, m \end{array}$$

Epigraph problem form The epigraph form of the standard problem (4.1) is the problem

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & f_0(x) - t \leq 0 \\ & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p \end{array} \tag{1.32}$$

with variables $x \in \mathbf{R}^n$ and $t \in \mathbf{R}$. We can easily see that it is equivalent to the original problem: (x, t) is optimal for (1.32) if and only if x is optimal for (1.22) and $t = f_0(x)$. Note that the objective function of the epigraph form problem is a linear function of the variables x, t.

The epigraph form problem (1.32) can be interpreted geometrically as an optimization problem in the 'graph space' (x, t) : we minimize t over the epigraph of f_0, subject to the constraints on x. This is illustrated in figure 4.1.

Implicit and explicit constraints

By a simple trick already mentioned in \$3.1.2, we can include any of the constraints implicitly in the objective function, by redefining its domain. As an extreme example, the standard form problem can be expressed as the unconstrained problem

$$\text{minimize} \quad F(x) \tag{1.33}$$

where we define the function F as f_0, but with domain restricted to the feasible set:

$$\mathbf{dom}\, F = \{x \in \mathbf{dom}\, f_0 \mid f_i(x) \leq 0, i = 1, \ldots, m, h_i(x) = 0, i = 1, \ldots, p\}$$

and $F(x) = f_0(x)$ for $x \in \mathbf{dom}\ F$. (Equivalently, we can define $F(x)$ to have value ∞ for x not feasible.) The problems (1.22) and (1.33) are clearly equivalent: they have the same feasible set, optimal points, and optimal value.

Of course this transformation is nothing more than a notational trick. Making the constraints implicit has not made the problem any easier to analyze or solve, even though the problem (4.12) is, at least nominally, unconstrained. In some ways the transformation makes the problem more difficult. Suppose, for example, that the objective f_0 in the original problem is differentiable, so in particular its domain is open. The restricted objective function F is probably not differentiable, since its domain is likely not to be open.

Conversely, we will encounter problems with implicit constraints, which we can then make explicit. As a simple example, consider the unconstrained problem

$$\text{minimize} \quad f(x) \tag{1.34}$$

where the function f is given by

$$f(x) = \begin{cases} x^T x & Ax = b \\ \infty & \text{otherwise} \end{cases}$$

Thus, the objective function is equal to the quadratic form $x^T x$ on the affine set defined by $Ax = b$, and ∞ off the affine set. since we can clearly restrict our attention to points that satisfy $Ax = b$, we say that the problem (1.34) has an implicit equality constraint $Ax = b$ hidden in the objective. We can make the implicit equality constraint explicit, by forming the equivalent problem

$$\begin{array}{ll} \text{minimize} & x^T x \\ \text{subject to} & Ax = b \end{array} \tag{1.35}$$

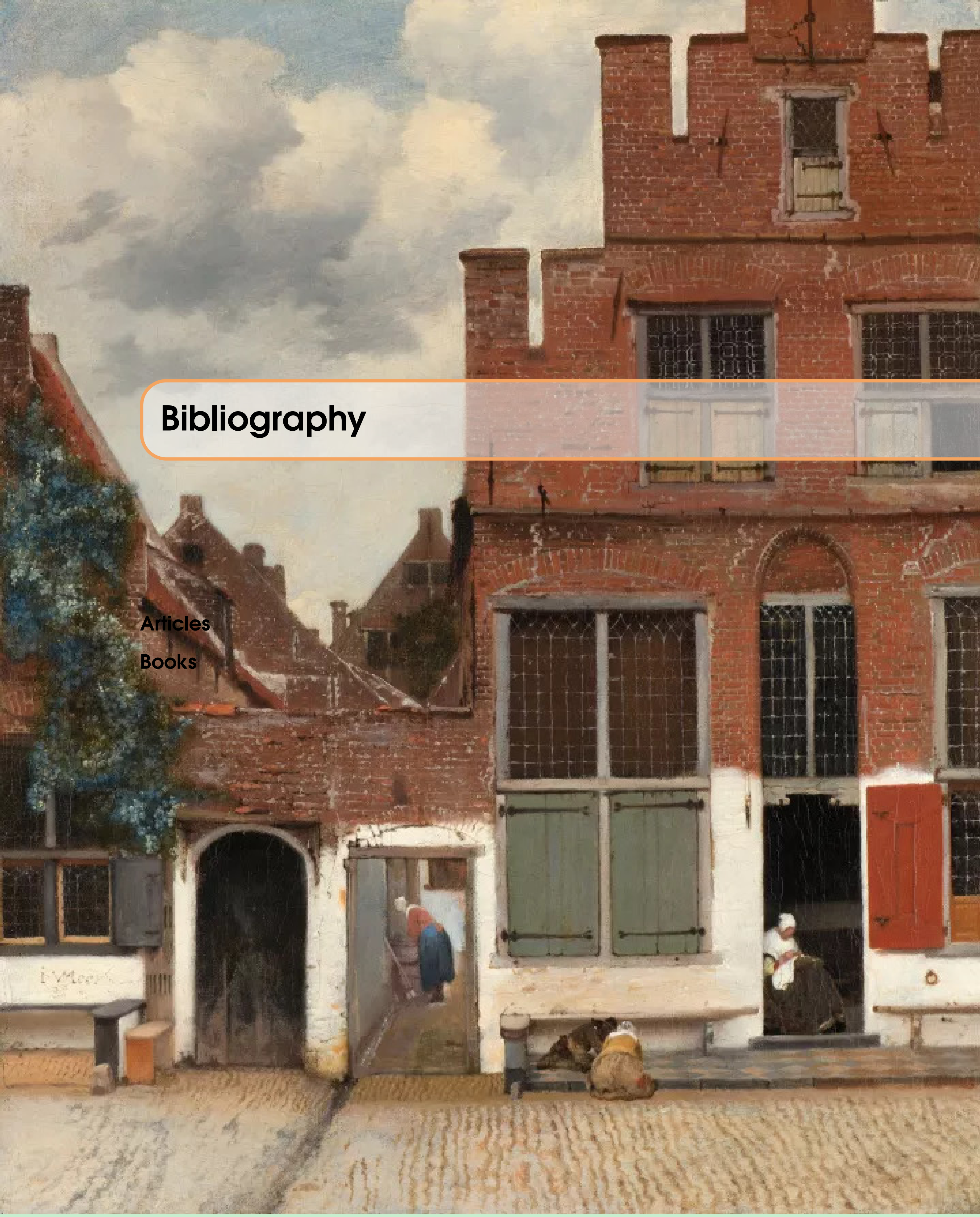
While the problems (1.34) and (1.35) are clearly equivalent, they are not the same. The problem (1.34) is unconstrained, but its objective function is not differentiable. The problem (1.35), however, has an equality constraint, but its objective and constraint functions are differentiable.

Parameter and oracle problem descriptions

For a problem in the standard form (1.22), there is still the question of how the objective and constraint functions are specified. In many cases these functions have some analytical or closed form, *i.e.*, are given by a formula or expression that involves the variable x as well as some parameters. Suppose, for example, the objective is quadratic, so it has the form $f_0(x) = (1/2)x^T Px + q^T x + r$. To specify the objective function we give the coefficients (also called problem parameters or problem data) $P \in \mathbf{S}^n, q \in \mathbf{R}^n$, and $r \in \mathbf{R}$. We call this a parameter problem description, since the specific problem to be solved (i.e., the problem instance) is specified by giving the values of the parameters that appear in the expressions for the objective and constraint functions. In other cases the objective and constraint functions are described by oracle models (which are also called black box or subroutine models). In an oracle model, we do not know f explicitly, but can evaluate

$f(x)$ (and usually also some derivatives) at any $x \in \operatorname{dom} f$. This is referred to as querying the oracle, and is usually associated with some cost, such as time. We are also given some prior information about the function, such as convexity and a bound on its values. As a concrete example of an oracle model, consider an unconstrained problem, in which we are to minimize the function f. The function value $f(x)$ and its gradient $\nabla f(x)$ are evaluated in a subroutine. We can call the subroutine at any $x \in \operatorname{dom} f$, but do not have access to its source code. Calling the subroutine with argument x yields (when the subroutine returns) $f(x)$ and $\nabla f(x)$. Note that in the oracle model, we never really know the function; we only know the function value (and some derivatives) at the points where we have queried the oracle. (We also know some given prior information about the function, such as differentiability and convexity.) In practice the distinction between a parameter and oracle problem description is not so sharp. If we are given a parameter problem description, we can construct an oracle for it, which simply evaluates the required functions and derivatives when queried. Most of the algorithms we study in part III work with an oracle model, but can be made more efficient when they are restricted to solve a specific parametrized family of problems.

Part Four



Bibliography