Chapter 1

Mathematical Foundations and Preliminaries

All numbers in this chapter are computed and checked in `scripts/ch01.py` (SymPy for exact values, NumPy, HiGHS and Clarabel for the cross-checks).

Exercise 1.1 Vector Norms

Compute $\|x\|_1$, $\|x\|_2$ and $\|x\|_\infty$ for $x = (3, -4, 2)^T$.

Apply the definitions entry by entry.

$$\|x\|_1 = |3| + |-4| + |2| = 9 \qquad \text{(sum of absolute values)}$$
$$\|x\|_2 = \sqrt{3^2 + (-4)^2 + 2^2} = \sqrt{29} \approx 5.3852 \qquad \text{(Euclidean length)}$$
$$\|x\|_\infty = \max\{3, 4, 2\} = 4 \qquad \text{(largest absolute entry)}$$

As expected, $\|x\|_\infty \le \|x\|_2 \le \|x\|_1$.

Exercise 1.2 Matrix Properties

For $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ find rank(A), null(A), and decide whether A is invertible.

(a) A is upper triangular, so its determinant is the product of the diagonal entries: $\det A = 1 \cdot 1 \cdot 1 = 1 \neq 0$. A square matrix with nonzero determinant has linearly independent columns, so $\operatorname{rank}(A) = 3$.

(b) Solve $Ax = 0$ by back substitution: the third row gives $x_3 = 0$; the second gives $x_2 + x_3 = 0$, so $x_2 = 0$; the first gives $x_1 + 2x_2 + x_3 = 0$, so $x_1 = 0$. Hence $\operatorname{null}(A) = \{0\}$. (This agrees with the rank–nullity theorem: $3 = \operatorname{rank}(A) + \dim \operatorname{null}(A) = 3 + 0$.)

(c) Yes, since $\det A \neq 0$. Back substitution on $AX = I$ gives

$$A^{-1} = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix},$$

and one checks $A^{-1}A = I$.

Exercise 1.3 Positive Definiteness

Classify $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ *as definite, semidefinite or indefinite.*

A symmetric matrix is positive definite if all its eigenvalues are positive, positive semidefinite if they are all ≥ 0, and indefinite if it has eigenvalues of both signs. For a symmetric 2×2 matrix $\begin{pmatrix} a & b \\ b & a \end{pmatrix}$ the eigenvalues are $a + b$ (eigenvector $(1, 1)$) and $a - b$ (eigenvector $(1, -1)$).

(a) Eigenvalues $1 + 2 = 3$ and $1 - 2 = -1$: one positive, one negative, so A is **indefinite**. For example $(1, -1)A(1, -1)^T = -2 < 0$ while $(1, 1)A(1, 1)^T = 6 > 0$.

(b) Eigenvalues 3 and 1, both positive: B is **positive definite**. (Also: leading minors $2 > 0$ and $\det B = 3 > 0$.)

(c) Eigenvalues 2 and 0: C is **positive semidefinite but not positive definite**, since $x^T Cx = (x_1 + x_2)^2 \geq 0$ with equality for $x = (1, -1)$.

Exercise 1.4 Gradient and Hessian

For $f(x, y) = x^3 + xy^2 - 3x$*, compute* ∇f *and* $\nabla^2 f$*, find and classify all critical points.*

(a) $\nabla f(x, y) = \begin{pmatrix} 3x^2 + y^2 - 3 \\ 2xy \end{pmatrix}$.

(b) $\nabla^2 f(x, y) = \begin{pmatrix} 6x & 2y \\ 2y & 2x \end{pmatrix}$.

(c) Critical points solve $3x^2 + y^2 = 3$ and $2xy = 0$. The second equation gives $x = 0$ or $y = 0$. If $x = 0$ then $y^2 = 3$, so $y = \pm\sqrt{3}$. If $y = 0$ then $x^2 = 1$, so $x = \pm 1$. The critical points are

$$(1, 0), \quad (-1, 0), \quad (0, \sqrt{3}), \quad (0, -\sqrt{3}).$$

(d) Use the second-derivative test (Chapter 1) on the Hessian at each point.

- $(1, 0)$: $\nabla^2 f = \mathrm{diag}(6, 2) \succ 0$, a **strict local minimum**, $f(1, 0) = -2$.
- $(-1, 0)$: $\nabla^2 f = \mathrm{diag}(-6, -2) \prec 0$, a **strict local maximum**, $f(-1, 0) = 2$.
- $(0, \pm\sqrt{3})$: $\nabla^2 f = \begin{pmatrix} 0 & \pm 2\sqrt{3} \\ \pm 2\sqrt{3} & 0 \end{pmatrix}$, with determinant $-12 < 0$, so one eigenvalue is positive and one negative: both are **saddle points** ($f = 0$ there).

Neither extremum is global: $f(x, 0) = x^3 - 3x \to -\infty$ as $x \to -\infty$ and $\to +\infty$ as $x \to +\infty$. This is typical of a nonconvex function.

Exercise 1.5 Optimization Problem Formulation

Products A and B use 2 h and 1 h of labour and 1 kg and 2 kg of raw material per item, with profits \$3 and \$2 per item; 100 h of labour and 80 kg of material are available. Formulate profit maximization.

Decisions. x_A, x_B: number of items of A and B to produce.

Problem.

$$\begin{array}{lll} \text{maximize} & 3x_A + 2x_B & \text{(profit, \$)} \\ \text{subject to} & 2x_A + x_B \le 100 & \text{(labour, hours)} \\ & x_A + 2x_B \le 80 & \text{(raw material, kg)} \\ & x_A,\ x_B \ge 0 & \text{(no negative production)} \end{array}$$

This is a linear program. If only whole items can be made, add $x_A, x_B \in \mathbb{Z}$; here this changes nothing, because the LP solution below is already integer.

Solution (not asked, but useful). The feasible region has vertices $(0,0)$, $(50,0)$, $(40,20)$ and $(0,40)$ with profits \$0, \$150, \$160 and \$80. The optimum is $x_A = 40$ items, $x_B = 20$ items, profit \$160, with both resources fully used. HiGHS and Clarabel return the same point and value. The shadow prices (dual values) are \$4/3 per hour of labour and \$1/3 per kg of material; they solve $2y_1 + y_2 = 3$, $y_1 + 2y_2 = 2$, and $100 \cdot \frac{4}{3} + 80 \cdot \frac{1}{3} = 160$ certifies optimality by strong duality.

Exercise 1.6 Taylor Approximation

Write the second-order Taylor approximation of $f(x,y) = e^{x+y}$ around $(0,0)$.

Every partial derivative of e^{x+y} of any order equals e^{x+y}, which is 1 at the origin. So $f(0,0) = 1$, $\nabla f(0,0) = (1,1)^T$ and $\nabla^2 f(0,0) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, and

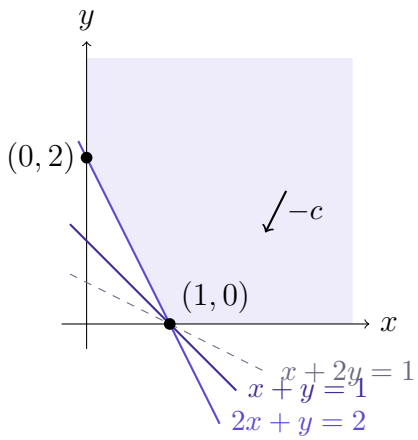
$$f(x,y) \approx 1 + (x+y) + \tfrac{1}{2}\,(x,y) \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 1 + x + y + \tfrac{1}{2}(x+y)^2 = 1 + x + y + \tfrac{1}{2}x^2 + xy + \tfrac{1}{2}y^2.$$

This matches $e^s \approx 1 + s + s^2/2$ with $s = x + y$. Check at $(0.1, 0.05)$: $e^{0.15} = 1.161834$, approximation 1.161250.

Exercise 1.7 Linear Program Geometry

Minimize $x + 2y$ subject to $x + y \ge 1$, $2x + y \ge 2$, $x, y \ge 0$: sketch, find the vertices, solve graphically.

(a) The feasible region lies above both lines $x + y = 1$ and $2x + y = 2$ in the first quadrant (shaded below). It is unbounded.



(b) A vertex is a feasible point where two independent constraints are active. Checking all pairs of the four boundary lines: $x = 0$ with $2x + y = 2$ gives $(0, 2)$ (feasible); $y = 0$ with either $x + y = 1$ or $2x + y = 2$ gives $(1, 0)$ (feasible; three constraints are active there); the remaining intersections $(0, 1)$ and $(0, 0)$ violate $2x + y \geq 2$. So the vertices are $(0, 2)$ and $(1, 0)$. Note that $x + y \geq 1$ is redundant: if $2x + y \geq 2$ and $x, y \geq 0$ then $x + y \geq 1$ (if $x \geq 1$ this is clear; if $x < 1$ then $x + y \geq x + 2 - 2x = 2 - x > 1$).

(c) The objective is bounded below on the region (it is ≥ 0 because $x, y \geq 0$), so an optimal vertex exists. Comparing vertices, $c^T(0, 2) = 4$ and $c^T(1, 0) = 1$. Graphically, slide the line $x + 2y = k$ down (direction $-c$) until it last touches the region: it touches only at $(1, 0)$. The optimum is $(x, y) = (1, 0)$ with value 1 (HiGHS and Clarabel agree).

Exercise 1.8 Quadratic Function Analysis

For $f(x) = \frac{1}{2}x^T Ax - b^T x$ with $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, $b = (1, 1)^T$: show f is convex, find its minimizer and minimum value.

(a) $\nabla^2 f(x) = A$ for every x. The eigenvalues of A are 3 and 1 (Exercise 1.3(b)), so $A \succ 0$ and f is (strictly) convex by the second-order condition.

(b) For a convex differentiable function, a point with zero gradient is a global minimizer. $\nabla f(x) = Ax - b = 0$ gives

$$x^\star = A^{-1}b = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix}.$$

It is the unique minimizer because f is strictly convex.

(c) Using $Ax^\star = b$: $f(x^\star) = \frac{1}{2}x^{\star T}b - b^T x^\star = -\frac{1}{2}b^T x^\star = -\frac{1}{2} \cdot \frac{2}{3} = -\frac{1}{3}$. CVXPY (Clarabel) returns the same point and value.