Solutions (LaTeX) to the uploaded optimization exam/assignments

Contents

1 Constrained Optimization (sheet, page 1)

Exercise (KKT on a square)**.** Let $\Omega = [0,4] \times [0,4]$ and

$$f(x) = \|x - (1,1)\|_2 , \qquad g = -f.$$

1) Minimum and maximum of f on Ω

Since f is the Euclidean distance to $(1,1)$:

- The minimum over Ω is attained at the projection of $(1,1)$ onto Ω, i.e. at $(1,1)$ itself:

$$\min_{\Omega} f = f(1,1) = 0.$$

- The maximum over Ω is attained at a farthest point of Ω from $(1,1)$, hence at a corner. Compute:

$$\begin{aligned} f(0,0) &= \sqrt{(1)^2 + (1)^2} = \sqrt{2}, \\ f(0,4) &= \sqrt{(1)^2 + (3)^2} = \sqrt{10}, \\ f(4,0) &= \sqrt{10}, \\ f(4,4) &= \sqrt{(3)^2 + (3)^2} = 3\sqrt{2}. \end{aligned}$$

  Thus $\max_{\Omega} f = 3\sqrt{2}$, attained at $(4,4)$.

2) Minimization as a linear-constrained problem and qualification

Write

$$\min_{x \in \mathbb{R}^2} f(x) \quad \text{s.t.} \quad \begin{cases} -x_1 \le 0, \\ x_1 - 4 \le 0, \\ -x_2 \le 0, \\ x_2 - 4 \le 0. \end{cases}$$

At any feasible point, the active constraint gradients are taken among $\{\pm e_1, \pm e_2\}$ and can never include both e_i and $-e_i$ for the same i simultaneously. Hence the active gradients are always linearly independent (LICQ holds everywhere on Ω).

3) No KKT point on the left edge $\{x_1 = 0\}$ and uniqueness of the KKT point in Ω

For $x \neq (1,1)$, f is differentiable with

$$\nabla f(x) = \frac{x - (1,1)}{\|x - (1,1)\|}.$$

Consider the minimization of f on Ω. On the left edge $x_1 = 0$, the constraint $c_1(x) = -x_1 \leq 0$ is active and $\nabla c_1 = -e_1$. If $x_2 \in (0,4)$, no other constraint is active, so KKT stationarity would read

$$\nabla f(x) + \mu_1 \nabla c_1 = 0 \quad \Longrightarrow \quad \nabla f(x) - \mu_1 e_1 = 0, \qquad \mu_1 \geq 0.$$

Taking the first component at $x_1 = 0$ gives

$$\frac{-1}{\|x - (1,1)\|} - \mu_1 = 0 \quad \Longrightarrow \quad \mu_1 = -\frac{1}{\|x - (1,1)\|} < 0,$$

a contradiction. If $x_2 \in \{0,4\}$, extra multipliers appear in the second coordinate, but the first coordinate equation is unchanged, and the same contradiction holds. Hence there is *no* KKT point on $\{x_1 = 0\}$.

By symmetry, the same argument excludes $x_1 = 4$, $x_2 = 0$, $x_2 = 4$ for the minimization problem. Therefore any KKT point must lie in the interior of Ω, where all multipliers vanish and $\nabla f(x) = 0$, which forces $x = (1,1)$ (the unique minimizer). Thus the only KKT point is $(1,1)$.

4) Maximization of f on Ω (equivalently minimization of $g = -f$)

We consider $\min_\Omega g(x)$ with $g = -f$. For $x \neq (1,1)$,

$$\nabla g(x) = -\nabla f(x) = -\frac{x - (1,1)}{\|x - (1,1)\|}.$$

KKT stationarity is $\nabla g(x) + \sum_{i=1}^4 \mu_i \nabla c_i(x) = 0$ with $\mu_i \geq 0$ and complementary slackness.

A quick way is to use the normal cone description: at a boundary point, $-\nabla g = \nabla f$ must belong to the normal cone of Ω. Concretely, one checks that KKT points for $\min g$ (i.e. $\max f$) are:

$$(4,4) \text{ (global maximizer)}, \quad (4,0), (0,4), (0,0) \text{ (corners)}, \quad (4,1), (0,1), (1,4), (1,0),$$

and, in a nonsmooth (subgradient) sense, $(1,1)$ as well (since $0 \in \partial f(1,1)$).

Nature of these points for $f|_\Omega$.

- $(1,1)$ is the global (hence strict) minimizer of f on Ω.
- $(4,4)$ is the global (hence strict) maximizer of f on Ω.
- The other listed boundary points are *saddle-type* for $f|_\Omega$: in any neighborhood inside Ω one can find points with larger and smaller distance to $(1,1)$ (e.g. move toward $(1,1)$ to decrease f, or toward $(4,4)$ / away from $(1,1)$ to increase f).

Exercise (Constraint qualification at a cusp)**.** Let

$$\Omega = \{(x_1,x_2) \in \mathbb{R}^2 : \ 0 \leq x_2 \leq x_1^2\}, \qquad f(x) = -x_2, \qquad x_\star = (0,0).$$

Write the constraints as

$$c_1(x) = -x_2 \leq 0, \qquad c_2(x) = x_2 - x_1^2 \leq 0.$$

1) Are constraints qualified at $x_\star$?

At $x_\star$, both constraints are active and

$$\nabla c_1(x_\star) = (0, -1), \qquad \nabla c_2(x_\star) = (0, 1).$$

These gradients are linearly dependent, so LICQ fails. Moreover, the Mangasarian–Fromovitz CQ fails: there is no direction d with $\langle \nabla c_1(x_\star), d \rangle < 0$ and $\langle \nabla c_2(x_\star), d \rangle < 0$ simultaneously (it would require $d_2 > 0$ and $d_2 < 0$).

2) Tangent directions vs. orthogonality to active gradients

The tangent (Bouligand) cone at $x_\star$ is

$$T_\Omega(x_\star) = \left\{ d \in \mathbb{R}^2 : \ \exists t_k \downarrow 0, \ x^{(k)} \in \Omega, \ \frac{x^{(k)} - x_\star}{t_k} \to d \right\}.$$

From $0 \le x_2^{(k)} \le (x_1^{(k)})^2$ we get

$$0 \le \frac{x_2^{(k)}}{t_k} \le t_k \left(\frac{x_1^{(k)}}{t_k} \right)^2 \to 0,$$

hence $d_2 = 0$. Conversely, any $(d_1, 0)$ is achievable by $x^{(k)} = (t_k d_1, 0)$, so

$$T_\Omega(x_\star) = \{(d_1, 0) : d_1 \in \mathbb{R}\}.$$

The active gradients span the x_2-axis; the space orthogonal to them is precisely the x_1-axis, i.e. $\{(d_1, 0)\}$, matching $T_\Omega(x_\star)$.

3) KKT at $x_\star$

$\nabla f(x) = (0, -1)$. KKT stationarity seeks $\mu_1, \mu_2 \ge 0$ such that

$$\nabla f(x_\star) + \mu_1 \nabla c_1(x_\star) + \mu_2 \nabla c_2(x_\star) = 0.$$

This becomes $(0, -1) + \mu_1(0, -1) + \mu_2(0, 1) = (0, 0)$, i.e. $\mu_2 = 1 + \mu_1$. Choosing $\mu_1 = 0$, $\mu_2 = 1$ yields stationarity, and complementary slackness holds since $c_1(x_\star) = c_2(x_\star) = 0$.

4) Is $x_\star$ a local minimizer?

No. Since $x_2 \ge 0$ on Ω, we have $f(x) = -x_2 \le 0 = f(x_\star)$ and any nearby feasible point with $x_2 > 0$ yields a strictly smaller objective value. Thus $x_\star$ is a (local) *maximizer*, not a minimizer.

Exercise (LP to standard form)**.** Given $c_1 \in \mathbb{R}^{n_1}$, $c_2 \in \mathbb{R}^{n_2}$, $l, u \in \mathbb{R}^{n_2}$, $A_1 \in \mathbb{R}^{p_1 \times n_1}$, $A_2 \in \mathbb{R}^{p_2 \times n_1}$, $B_2 \in \mathbb{R}^{p_2 \times n_2}$, $b_1 \in \mathbb{R}^{p_1}$, $b_2 \in \mathbb{R}^{p_2}$, consider

$$\max \ c_1^\top x_1 + c_2^\top x_2 \quad \text{s.t.} \quad A_1 x_1 = b_1, \quad A_2 x_1 + B_2 x_2 \le b_2, \quad l \le x_2 \le u,$$

with x_1 free.

Step 1: split the free variable. Write $x_1 = x_1^+ - x_1^-$ with $x_1^+, x_1^- \ge 0$.

Step 2: shift the lower bound on x_2. Let $y = x_2 - l$, so $y \ge 0$ and $y \le u - l$ (assume $u \ge l$ componentwise).

Step 3: convert inequalities to equalities with slack variables.

- For $A_2x_1 + B_2x_2 \le b_2$, substitute $x_1 = x_1^+ - x_1^-$ and $x_2 = l + y$:

$$A_2x_1^+ - A_2x_1^- + B_2y \le b_2 - B_2l.$$

  Add slack $s \ge 0$ in $\mathbb{R}^{p_2}$:

$$A_2x_1^+ - A_2x_1^- + B_2y + s = b_2 - B_2l.$$

- For the upper bound $y \le u - l$, add slack $t \ge 0$ in $\mathbb{R}^{n_2}$:

$$y + t = u - l.$$

Final standard-form system. Define the nonnegative decision vector

$$x = \begin{pmatrix} x_1^+ \\ x_1^- \\ y \\ s \\ t \end{pmatrix} \ge 0.$$

Then the constraints become $Ax = b$ with

$$A = \begin{pmatrix} A_1 & -A_1 & 0 & 0 & 0 \\ A_2 & -A_2 & B_2 & I_{p_2} & 0 \\ 0 & 0 & I_{n_2} & 0 & I_{n_2} \end{pmatrix}, \qquad b = \begin{pmatrix} b_1 \\ b_2 - B_2l \\ u - l \end{pmatrix}.$$

The objective becomes

$$\max \underbrace{\left(c_1^\top x_1^+ - c_1^\top x_1^- + c_2^\top y\right)}_{\text{linear in } x} + c_2^\top l,$$

where the constant $c_2^\top l$ can be dropped if desired.

2 Gradient and convexity (sheet, page 2)

Exercise (Taylor / integral form)**.** Let $f \in C^2(\Omega)$ on a convex open set $\Omega \subset \mathbb{R}^n$. Show that for all $x, y \in \Omega$,

$$\nabla f(y) - \nabla f(x) = \int_0^1 \nabla^2 f((1-t)x + ty)\,(y - x)\,dt.$$

Proof. Fix $p \in \mathbb{R}^n$ and define the scalar function

$$\phi(t) = \langle p, \nabla f((1-t)x + ty)\rangle.$$

By the chain rule,

$$\phi'(t) = \left\langle p, \nabla^2 f((1-t)x + ty)\,(y - x)\right\rangle.$$

Integrating from 0 to 1 gives

$$\langle p, \nabla f(y) - \nabla f(x)\rangle = \phi(1) - \phi(0) = \int_0^1 \left\langle p, \nabla^2 f((1-t)x + ty)\,(y - x)\right\rangle dt.$$

Since this holds for all p, the claimed vector identity follows. □

Exercise (Convexity of a quadratic form)**.** Let A be positive (semi-)definite. Show that $q(x) = x^\top Ax$ is convex.

Proof (by definition). For $t \in [0,1]$ and $x, y \in \mathbb{R}^n$,

$$\begin{aligned} q(tx + (1-t)y) &= (tx + (1-t)y)^\top A(tx + (1-t)y) \\ &= tx^\top Ax + (1-t)y^\top Ay - t(1-t)(x-y)^\top A(x-y) \\ &\le tq(x) + (1-t)q(y), \end{aligned}$$

because $(x-y)^\top A(x-y) \ge 0$ when $A \succeq 0$. Hence q is convex. □

Exercise (Isolated local minimizers are strict)**.** If $f : \Omega \to \mathbb{R}$ has an isolated local minimizer at $x \in \Omega$, prove that it is a strict local minimizer.

Proof. Since x is a local minimizer, there exists $r > 0$ such that $f(z) \ge f(x)$ for all $z \in B(x,r) \cap \Omega$. If x were not strict, then for every k there would exist $z_k \in (B(x, 1/k) \cap \Omega) \setminus \{x\}$ with $f(z_k) = f(x)$. Then $z_k \to x$ and $f(z_k) = f(x)$, so x would not be isolated among local minimizers/points with minimal value in a neighborhood, a contradiction. □

Exercise (Stationary points of two functions)**.** Consider

$$f(x_1, x_2) = 8x_1 + 12x_2 + x_1^2 - 2x_2^2, \qquad g(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

(a) f: unique stationary point, not a local extremum

Compute

$$\nabla f(x_1, x_2) = \begin{pmatrix} 2x_1 + 8 \\ -4x_2 + 12 \end{pmatrix}.$$

Thus $\nabla f = 0$ iff $x_1 = -4$ and $x_2 = 3$, so the stationary point is unique: $x^\star = (-4, 3)$.

The Hessian is constant:

$$\nabla^2 f = \begin{pmatrix} 2 & 0 \\ 0 & -4 \end{pmatrix},$$

which is indefinite. Hence $x^\star$ is a saddle point (neither a local minimum nor a local maximum).

(b) g: unique stationary point, global minimum

Compute

$$\nabla g(x_1, x_2) = \begin{pmatrix} -400x_1(x_2 - x_1^2) - 2(1 - x_1) \\ 200(x_2 - x_1^2) \end{pmatrix}.$$

Setting the second component to zero yields $x_2 = x_1^2$. Plugging into the first gives $-2(1-x_1) = 0$, hence $x_1 = 1$ and $x_2 = 1$. So the stationary point is unique: $(1, 1)$.

Moreover, g is a sum of squares, so $g \ge 0$, and $g(x) = 0$ iff $x_2 = x_1^2$ and $x_1 = 1$, i.e. only at $(1, 1)$. Therefore $(1, 1)$ is the unique global minimizer.

Exercise (Rates of convergence)**.** Let $x^{(k)} \to x_\star$ and assume $\exists M > 0, \exists p > 0$ such that

$$\frac{\left\|x^{(k+1)} - x_\star\right\|}{\left\|x^{(k)} - x_\star\right\|^p} \le M \quad \forall k.$$

1) Q-quadratic ⇒ Q-linear

If $p = 2$, then $\|e_{k+1}\| \le M \|e_k\|^2$ with $e_k = x^{(k)} - x_\star$. Since $e_k \to 0$, there exists K such that $\|e_k\| \le 1$ for all $k \ge K$, hence

$$\|e_{k+1}\| \le M \|e_k\| \qquad (k \ge K),$$

which is Q-linear (order $p = 1$) from that point onward.

2) Order of convergence of three sequences

Let the limit be 1.

- $x_k = 1 + (1/3)^k$: $e_k = (1/3)^k$ and $e_{k+1} = (1/3)e_k$, so Q-linear (order $p = 1$).
- $x_k = 1 + e^{-e^k}$: $e_k = \exp(-e^k)$ and $e_{k+1} = \exp(-e^{k+1}) = \exp(-e \cdot e^k) = \big(\exp(-e^k)\big)^e = e_k^e$, so the order is $p = e$ (superlinear).
- $x_k = 1 + \frac{1}{k^2}$: $e_k = 1/k^2$ and $e_{k+1}/e_k \to 1$. The convergence is sublinear (certainly not geometric with ratio < 1). In the given definition, one can take $p = 1$ with $M = 1$ since $e_{k+1} \le e_k$, but the rate is not Q-linear in the common (geometric) sense.

3) Correct digits

If $\|e_k\| \approx 10^{-P}$, then:

- Q-linear: $\|e_{k+1}\| \le M\,\|e_k\|$ implies

$$-\log_{10}\|e_{k+1}\| \ \gtrsim\ P - \log_{10} M,$$

  so the number of correct digits increases by an *additive constant* per iteration (asymptotically).

- Q-quadratic: $\|e_{k+1}\| \le M\,\|e_k\|^2$ implies

$$-\log_{10}\|e_{k+1}\| \ \gtrsim\ 2P - \log_{10} M,$$

  so the number of correct digits is (approximately) *doubled* each iteration (up to an additive constant).

Exercise (Descent direction and 1D minimization along a ray)**.** Let $f(x_1, x_2) = (x_1 + x_2^2)^2$, $x = (0, 1)$ and $p = (-1, 1)$.

Descent check

Let $s = x_1 + x_2^2$. Then $\nabla f(x) = (2s,\ 4x_2 s)$. At $x = (0, 1)$, $s = 1$ and $\nabla f(x) = (2, 4)$, hence

$$\langle \nabla f(x), p\rangle = (2, 4) \cdot (-1, 1) = 2 > 0.$$

Therefore p is *not* a descent direction for minimizing f at x (the descent direction is $-p$).

Minimization of f along $\{x + tp : t \ge 0\}$

Along $x(t) = x + tp = (-t, 1 + t)$,

$$f(x(t)) = \big(-t + (1 + t)^2\big)^2 = (1 + t + t^2)^2.$$

The function $h(t) = (1 + t + t^2)^2$ satisfies

$$h'(t) = 2(1 + t + t^2)(1 + 2t) > 0 \qquad \forall t \ge 0,$$

so h is strictly increasing on $[0, \infty)$. Hence the unique minimizer on the ray is $t_\star = 0$, i.e. the point x itself.

3 Homework assignment (sheet, page 3)

Exercise (Kantorovich inequality (one proof path))**.** Let $A \succ 0$ be symmetric with eigenvalues $0 < \lambda_1 \le \cdots \le \lambda_n$, and $x \in \mathbb{R}^n$. Show

$$\|x\|_2^2 \le \langle x, Ax\rangle^{1/2} \left\langle x, A^{-1}x\right\rangle^{1/2} \le \frac{\lambda_1 + \lambda_n}{2\sqrt{\lambda_1\lambda_n}} \|x\|_2^2 .$$

Lower bound

Let $u = A^{1/2}x$ and $v = A^{-1/2}x$. Then $\langle u, v\rangle = \|x\|_2^2$ and by Cauchy–Schwarz,

$$\|x\|_2^4 = \langle u, v\rangle^2 \le \|u\|_2^2 \|v\|_2^2 = \langle x, Ax\rangle \left\langle x, A^{-1}x\right\rangle .$$

Upper bound (following the hints)

Diagonalize $A = Q^\top \Lambda Q$ with $\Lambda = \operatorname{diag}(\lambda_i)$ and set $y = Qx$ (so $\|y\| = \|x\|$). Then

$$\langle x, Ax\rangle = \sum_{i=1}^n \lambda_i y_i^2, \qquad \left\langle x, A^{-1}x\right\rangle = \sum_{i=1}^n \lambda_i^{-1} y_i^2.$$

Thus it suffices to prove the inequality for diagonal A.

By scaling, one may assume $\lambda_n = 1/\lambda_1$ (normalize by the geometric mean of the extreme eigenvalues). Using $ab \le \frac{1}{2}(a^2 + b^2)$ with $a = \sqrt{\lambda_i}\,|y_i|$ and $b = \lambda_i^{-1/2}|y_i|$ yields

$$\sqrt{\sum_i \lambda_i y_i^2}\,\sqrt{\sum_i \lambda_i^{-1} y_i^2} \le \frac{1}{2}\sum_i (\lambda_i + \lambda_i^{-1}) y_i^2.$$

Since $\lambda_1 \le \lambda_i \le \lambda_n$ implies $\lambda_i + \lambda_i^{-1} \le \lambda_1 + \lambda_1^{-1}$ (the function $t + 1/t$ increases on $[1, \infty)$ and is symmetric), we get

$$\frac{1}{2}\sum_i (\lambda_i + \lambda_i^{-1}) y_i^2 \le \frac{1}{2}\left(\lambda_1 + \frac{1}{\lambda_1}\right)\sum_i y_i^2 = \frac{\lambda_1 + \lambda_n}{2\sqrt{\lambda_1\lambda_n}} \|x\|_2^2 ,$$

which is the desired upper bound. □

Exercise (Steepest descent for a quadratic and condition number)**.** Assume $A \succ 0$ and consider

$$q(x) = \frac{1}{2}\langle x, Ax\rangle - \langle b, x\rangle .$$

Then $\nabla q(x) = Ax - b$. Steepest descent uses $p^{(k)} = -(\nabla q(x^{(k)})) = b - Ax^{(k)} =: r^{(k)}$ and chooses the optimal step size

$$\alpha_k = \arg\min_{\alpha \ge 0} q(x^{(k)} + \alpha r^{(k)}) = \frac{\left\langle r^{(k)}, r^{(k)}\right\rangle}{\left\langle r^{(k)}, Ar^{(k)}\right\rangle}.$$

Let $x_\star$ be the minimizer ($Ax_\star = b$), $e^{(k)} = x^{(k)} - x_\star$, and note $r^{(k)} = b - Ax^{(k)} = -Ae^{(k)}$.

2) Energy decrease identity

Using $e^{(k+1)} = e^{(k)} + \alpha_k r^{(k)}$ and $r^{(k)} = -Ae^{(k)}$,

$$\begin{aligned}\left\|e^{(k+1)}\right\|_A^2 &= \left\langle e^{(k)} + \alpha_k r^{(k)}, A(e^{(k)} + \alpha_k r^{(k)})\right\rangle \\ &= \left\|e^{(k)}\right\|_A^2 + 2\alpha_k \left\langle e^{(k)}, Ar^{(k)}\right\rangle + \alpha_k^2 \left\langle r^{(k)}, Ar^{(k)}\right\rangle .\end{aligned}$$

But $\left\langle e^{(k)}, Ar^{(k)} \right\rangle = \left\langle e^{(k)}, A(b - Ax^{(k)}) \right\rangle = \left\langle e^{(k)}, -A^2 e^{(k)} \right\rangle = -\left\langle r^{(k)}, r^{(k)} \right\rangle$. Plugging this and the expression of α_k gives

$$\left\| e^{(k+1)} \right\|_A^2 = \left\| e^{(k)} \right\|_A^2 - \frac{\left\| r^{(k)} \right\|_2^4}{\left\langle r^{(k)}, Ar^{(k)} \right\rangle}.$$

3) Linear rate via Kantorovich

Apply the Kantorovich inequality to $r^{(k)}$:

$$\frac{\left\| r^{(k)} \right\|_2^4}{\left\langle r^{(k)}, Ar^{(k)} \right\rangle} \geq \frac{4\lambda_1 \lambda_n}{(\lambda_1 + \lambda_n)^2} \left\langle r^{(k)}, A^{-1} r^{(k)} \right\rangle.$$

Since $r^{(k)} = -Ae^{(k)}$, we have $\left\langle r^{(k)}, A^{-1} r^{(k)} \right\rangle = \left\langle e^{(k)}, Ae^{(k)} \right\rangle = \left\| e^{(k)} \right\|_A^2$. Therefore

$$\left\| e^{(k+1)} \right\|_A^2 \leq \left(1 - \frac{4\lambda_1 \lambda_n}{(\lambda_1 + \lambda_n)^2} \right) \left\| e^{(k)} \right\|_A^2 = \left(\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1} \right)^2 \left\| e^{(k)} \right\|_A^2,$$

hence

$$\left\| e^{(k+1)} \right\|_A \leq \frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1} \left\| e^{(k)} \right\|_A.$$

4) Condition number form

Let $\kappa_2(A) = \lambda_n / \lambda_1 = \|A\|_2 \left\| A^{-1} \right\|_2$. Then

$$\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1} = \frac{\kappa_2(A) - 1}{\kappa_2(A) + 1},$$

so

$$\left\| e^{(k)} \right\|_A \leq \left(\frac{\kappa_2(A) - 1}{\kappa_2(A) + 1} \right)^k \left\| e^{(0)} \right\|_A.$$

5) Comment

The convergence is Q-linear (geometric). The factor deteriorates as $\kappa_2(A)$ grows: ill-conditioning makes the level sets highly elongated, so steepest descent “zig-zags” and progresses slowly.

Exercise (Code exercise on Rosenbrock: one reproducible reference run)**.** We report one reference run for the Rosenbrock function

$$g(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2, \qquad x_\star = (1, 1).$$

We used (i) steepest descent with Armijo backtracking line search (initial step 1, contraction 1/2, Armijo parameter 10^{-4}), and (ii) Newton's method with the same line search for globalization.

The iteration counts to reach $\left\| x^{(k)} - x_\star \right\| \leq \varepsilon$ were:

Start $x^{(0)}$	Method	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-6}$
$(1.2, 1.2)$	Steepest descent	5189	13708
$(1.2, 1.2)$	Newton	6	7
$(-1.2, 1)$	Steepest descent	5932	14446
$(-1.2, 1)$	Newton	20	21

Remark. Exact numbers depend on the precise line-search and stopping rules; however, the qualitative behavior (very slow gradient descent, fast Newton near the minimizer) is robust.

4 Final Exam (pages 4–5)

Conventions

In this section, $\|\cdot\|$ is the Euclidean norm, and $\langle\cdot,\cdot\rangle$ the standard inner product. Matrices $B \in \mathbb{R}^{d\times d}$ are symmetric, with eigenvalues $\lambda_1 \le \cdots \le \lambda_d$ and an orthonormal eigenbasis (v_i).

Exercise (PSD characterizations and a second-order condition). **1)** $B \succeq 0 \iff \lambda_1 \ge 0$

If $B \succeq 0$, then for every eigenvector v_i with $\|v_i\| = 1$,

$$0 \le \langle v_i, Bv_i\rangle = \langle v_i, \lambda_i v_i\rangle = \lambda_i,$$

so all $\lambda_i \ge 0$ and in particular $\lambda_1 \ge 0$.

Conversely, if $\lambda_1 \ge 0$ then all $\lambda_i \ge \lambda_1 \ge 0$. For any $v = \sum_i \alpha_i v_i$,

$$\langle v, Bv\rangle = \sum_i \lambda_i \alpha_i^2 \ge 0,$$

hence $B \succeq 0$.

2) Testing only on a half-space

Let $v_0 \neq 0$. We show:

$$B \succeq 0 \quad\iff\quad \forall v \in \mathbb{R}^d,\ \langle v, v_0\rangle > 0 \ \Rightarrow\ \langle v, Bv\rangle \ge 0.$$

The forward implication is immediate.

For the reverse implication, assume the half-space condition holds and take any $w \in \mathbb{R}^d\setminus\{0\}$. If $\langle w, v_0\rangle > 0$, apply the assumption to $v = w$. If $\langle w, v_0\rangle < 0$, apply the assumption to $v = -w$ to get $\langle -w, B(-w)\rangle = \langle w, Bw\rangle \ge 0$. If $\langle w, v_0\rangle = 0$, pick any u with $\langle u, v_0\rangle > 0$ and consider $v_\varepsilon = w + \varepsilon u$. Then $\langle v_\varepsilon, v_0\rangle > 0$ for all $\varepsilon > 0$, hence $\langle v_\varepsilon, Bv_\varepsilon\rangle \ge 0$. Letting $\varepsilon \downarrow 0$ and using continuity of $v \mapsto \langle v, Bv\rangle$ yields $\langle w, Bw\rangle \ge 0$. Thus $B \succeq 0$.

3) Half-space minimality $\Rightarrow$ PSD Hessian

Let $f \in C^2(\mathbb{R}^d)$, x_0 be a critical point ($\nabla f(x_0) = 0$), let $v_0 \in \mathbb{R}^d$, and set

$$H = \{x \in \mathbb{R}^d :\ \langle x, v_0\rangle \ge \langle x_0, v_0\rangle\}.$$

Assume $f(x) \ge f(x_0)$ for all $x \in H$. Define $\varphi(t) = f(x_0 + tv)$ for any v with $\langle v, v_0\rangle > 0$. Then for all sufficiently small $t \ge 0$, $x_0 + tv \in H$, hence $\varphi(t) \ge \varphi(0)$. Therefore $\varphi'(0) = 0$ and $\varphi''(0) \ge 0$. Since $\varphi'(0) = \langle \nabla f(x_0), v\rangle = 0$, we have

$$\varphi''(0) = v^\top \nabla^2 f(x_0)\, v \ge 0 \qquad \text{for all } v \text{ with } \langle v, v_0\rangle > 0.$$

By part (2) applied to $B = \nabla^2 f(x_0)$, it follows that $\nabla^2 f(x_0) \succeq 0$. □

Exercise (Trust-region problem). Let B be symmetric, $b \in \mathbb{R}^d$, and

$$f(x) = \frac{1}{2}\langle x, Bx\rangle - \langle b, x\rangle.$$

We consider the trust-region problem

$$\min f(x) \quad \text{s.t.} \quad \|x\| \le \delta. \tag{TR}$$

1) Gradient and Hessian

Since B is symmetric,

$$\nabla f(x) = Bx - b, \qquad \nabla^2 f(x) = B.$$

2) Existence/uniqueness of an unconstrained minimizer on $\mathbb{R}^d$

Write $b = \sum_i \beta_i v_i$. Then

$$f\Big(\sum_i \alpha_i v_i\Big) = \frac{1}{2}\sum_i \lambda_i \alpha_i^2 - \sum_i \beta_i \alpha_i.$$

- If $\lambda_1 > 0$ (i.e. $B \succ 0$), f is strongly convex and has a unique global minimizer $x_\star = B^{-1}b$.
- If $\lambda_1 = 0$ (i.e. $B \succeq 0$ but singular), f is convex. It is bounded below iff $b \in \mathrm{Range}(B)$ (equivalently $\beta_i = 0$ for all $\lambda_i = 0$). In that case, minimizers exist but are not unique: any solution of $Bx = b$ is a minimizer, and one can add any vector in $\ker(B)$. If $b \notin \mathrm{Range}(B)$, then f is unbounded below along directions in $\ker(B)$.
- If $\lambda_1 < 0$ (indefinite), then f is unbounded below on $\mathbb{R}^d$ (move along an eigenvector of a negative eigenvalue).

3) f is bounded on B_δ and attains its minimum

The ball $B_\delta = \{x : \|x\| \le \delta\}$ is compact and f is continuous, hence f is bounded and attains its minimum on B_δ (Weierstrass theorem).

4) Smooth constraint function

Take

$$g_\delta(x) = \delta^2 - \|x\|^2.$$

Then $g_\delta \in C^\infty(\mathbb{R}^d)$ and $\|x\| \le \delta \iff g_\delta(x) \ge 0 \iff h(x) := \|x\|^2 - \delta^2 \le 0$. Thus (TR) is

$$\min f(x) \quad \text{s.t.} \quad h(x) \le 0, \qquad h(x) = \|x\|^2 - \delta^2.$$

5) Constraint qualification at a solution

Let $x_\star$ solve (TR). If $\|x_\star\| < \delta$ the constraint is inactive, so qualification is automatic. If $\|x_\star\| = \delta$, then $\nabla h(x_\star) = 2x_\star \neq 0$ (since $\delta > 0$), hence LICQ holds.

6) KKT conditions

The Lagrangian is $L(x,\lambda) = f(x) + \frac{\lambda}{2}(\|x\|^2 - \delta^2)$. KKT yields $\lambda_\star \ge 0$ such that

$$\begin{cases} \text{(stationarity)} & \nabla f(x_\star) + \lambda_\star x_\star = 0 \iff Bx_\star - b = -\lambda_\star x_\star, \\ \text{(primal feasibility)} & \|x_\star\| \le \delta, \\ \text{(complementary slackness)} & \lambda_\star(\|x_\star\|^2 - \delta^2) = 0. \end{cases}$$

7) Identity (2)

For any $x \in \mathbb{R}^d$,

$$\begin{aligned} f(x) - f(x_\star) &= \frac{1}{2} \langle x, Bx \rangle - \langle b, x \rangle - \Big(\frac{1}{2} \langle x_\star, Bx_\star \rangle - \langle b, x_\star \rangle \Big) \\ &= \frac{1}{2} \langle x - x_\star, B(x - x_\star) \rangle + \langle Bx_\star - b, x - x_\star \rangle . \end{aligned}$$

Using stationarity $Bx_\star - b = -\lambda_\star x_\star$ gives

$$f(x) - f(x_\star) = \frac{1}{2} \langle x - x_\star, B(x - x_\star) \rangle - \lambda_\star \langle x_\star, x - x_\star \rangle .$$

Now expand $\|x\|^2 - \|x_\star\|^2 = \|x - x_\star\|^2 + 2 \langle x_\star, x - x_\star \rangle$ to obtain

$$\frac{1}{2} \langle x - x_\star, (B + \lambda_\star I)(x - x_\star) \rangle = f(x) - f(x_\star) + \frac{\lambda_\star}{2} \big(\|x\|^2 - \|x_\star\|^2 \big),$$

which is exactly (2).

8) If $\lambda_\star = 0$, then $B \succeq 0$

If $\lambda_\star = 0$, (2) becomes

$$\frac{1}{2} \langle x - x_\star, B(x - x_\star) \rangle = f(x) - f(x_\star).$$

Since $x_\star$ minimizes f on B_δ, the right-hand side is ≥ 0 for all $x \in B_\delta$. Hence $\langle v, Bv \rangle \geq 0$ for all v of the form $v = x - x_\star$ with $x \in B_\delta$.

Now fix any $w \neq 0$ and choose $t > 0$ small enough so that $x = x_\star + tw \in B_\delta$ (possible since B_δ has nonempty interior). Then $v = tw$ and $\langle w, Bw \rangle = \frac{1}{t^2} \langle v, Bv \rangle \geq 0$. Thus $B \succeq 0$.

9) $B + \lambda_\star I \succeq 0$ in all cases

If $\lambda_\star = 0$, this is immediate from (8).

Assume $\lambda_\star > 0$. Then complementary slackness forces $\|x_\star\| = \delta$. For any x on the sphere $\|x\| = \delta$, (2) reduces to

$$\frac{1}{2} \langle x - x_\star, (B + \lambda_\star I)(x - x_\star) \rangle = f(x) - f(x_\star) \geq 0,$$

so the quadratic form of $B + \lambda_\star I$ is nonnegative on all vectors $v = x - x_\star$ with $\|x\| = \delta$. These v satisfy $\langle v, -x_\star \rangle \geq 0$ (by Cauchy–Schwarz: $\langle x, x_\star \rangle \leq \|x\| \, \|x_\star\| = \delta^2$). Moreover, for any w with $\langle w, -x_\star \rangle > 0$ one can choose $\alpha > 0$ such that $x = x_\star + \alpha w$ lies on the sphere $\|x\| = \delta$, hence w is a positive multiple of some $x - x_\star$. Therefore $\langle w, (B + \lambda_\star I)w \rangle \geq 0$ for all w with $\langle w, -x_\star \rangle > 0$. By Exercise 1(2), this implies $B + \lambda_\star I \succeq 0$.

10) Why is $x(\lambda) = (B + \lambda I)^{-1} b$ well-defined for $\lambda > -\lambda_1$?

If $\lambda > -\lambda_1$, then all eigenvalues of $B + \lambda I$ are $\lambda_i + \lambda \geq \lambda_1 + \lambda > 0$, hence $B + \lambda I$ is positive definite and invertible.

11) $\psi(\lambda) = \|x(\lambda)\|^2$ is decreasing on $(-\lambda_1, \infty)$

Write $b = \sum_i \beta_i v_i$. Then

$$x(\lambda) = \sum_{i=1}^{d} \frac{\beta_i}{\lambda_i + \lambda} v_i, \qquad \psi(\lambda) = \sum_{i=1}^{d} \frac{\beta_i^2}{(\lambda_i + \lambda)^2}.$$

Each term is strictly decreasing in λ on $(-\lambda_i, \infty)$, hence ψ is decreasing on $(-\lambda_1, \infty)$.

12) Uniqueness of the solution to $\psi(\lambda) = \delta^2$

Assume $\langle b, v_1 \rangle = \beta_1 \neq 0$ and $\lambda_1 < 0$. Then as $\lambda \downarrow -\lambda_1$,

$$\psi(\lambda) \geq \frac{\beta_1^2}{(\lambda_1 + \lambda)^2} \to +\infty,$$

while $\psi(\lambda) \to 0$ as $\lambda \to +\infty$. By continuity and strict monotonicity, the equation $\psi(\lambda) = \delta^2$ has a unique solution $\hat{\lambda} \in (-\lambda_1, \infty)$.

13) $x(\hat{\lambda})$ solves (TR)

By construction, $x(\hat{\lambda})$ satisfies $(B + \hat{\lambda} I) x(\hat{\lambda}) = b$ and $\left\| x(\hat{\lambda}) \right\| = \delta$. Thus the KKT system in (6) holds with $\lambda_\star = \hat{\lambda} > 0$. From (9), $B + \hat{\lambda} I \succeq 0$, and the identity (2) implies for all x with $\|x\| \leq \delta$,

$$f(x) - f(x(\hat{\lambda})) = \frac{1}{2} \left\langle x - x(\hat{\lambda}), (B + \hat{\lambda} I)(x - x(\hat{\lambda})) \right\rangle - \frac{\hat{\lambda}}{2} \left(\|x\|^2 - \delta^2 \right) \geq 0,$$

since the first term is ≥ 0 and the second term is also ≥ 0 (because $\hat{\lambda} > 0$ and $\|x\| \leq \delta$). Hence $x(\hat{\lambda})$ is a global minimizer of (TR).

14) Algorithm to find $\hat{\lambda}$

Because ψ is continuous and strictly decreasing, one may solve $\psi(\lambda) - \delta^2 = 0$ by:

- **Bisection** on an interval $[\lambda_L, \lambda_U] \subset (-\lambda_1, \infty)$ with $\psi(\lambda_L) \geq \delta^2 \geq \psi(\lambda_U)$ (robust), or
- **Safeguarded Newton** using $\psi'(\lambda) = -2 \sum_i \beta_i^2 / (\lambda_i + \lambda)^3$ (faster).

Exercise (Discrete optimal transport as an LP)**.** Let $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R}^d$. A transport plan is $(x_{ij})_{1 \leq i,j \leq n}$ such that

$$x_{ij} \geq 0, \qquad \sum_{j=1}^{n} x_{ij} = 1 \quad (\forall i), \qquad \sum_{i=1}^{n} x_{ij} = 1 \quad (\forall j).$$

1) Meaning of the constraints

x_{ij} is the mass moved from a_i to b_j. Row sums equal 1 mean each a_i sends all its unit mass. Column sums equal 1 mean each b_j receives exactly one unit mass. Nonnegativity means no "negative mass".

2) Linearity of the L^p optimal transport problem

For $p \geq 1$ define costs $c_{ij} = \|a_i - b_j\|^p$ and objective

$$F(x) = \sum_{i,j} c_{ij} x_{ij}.$$

F is linear in the variables x_{ij} and the constraints are linear, so this is a linear program.

3) Existence of an optimal plan

The feasible set is the Birkhoff polytope (doubly stochastic matrices), which is nonempty, closed, and bounded; hence compact. Since F is continuous, a minimizer exists.

4) Dual variables and complementary slackness

Write the primal in standard form:

$$\min_{x\geq 0} \sum_{i,j} c_{ij}x_{ij} \text{ s.t. } \sum_j x_{ij} = 1, \; \sum_i x_{ij} = 1.$$

The dual is

$$\max_{\alpha,\beta} \sum_i \alpha_i + \sum_j \beta_j \quad \text{s.t.} \quad \alpha_i + \beta_j \leq c_{ij} \;\; \forall i,j.$$

Set $\gamma_{ij} = c_{ij} - \alpha_i - \beta_j \geq 0$. Then

$$c_{ij} = \gamma_{ij} + \alpha_i + \beta_j, \qquad \gamma_{ij} \geq 0, \qquad \gamma_{ij}x_{ij} = 0 \;\; \text{(complementary slackness)}.$$

This is exactly the requested form, with the additional condition $\gamma_{ij}x_{ij} = 0$.

5) The 2×2 case, $p = 2$

For $n = 2$, feasibility forces

$$x = \begin{pmatrix} t & 1-t \\ 1-t & t \end{pmatrix}, \qquad t \in [0,1].$$

Let $c_{ij} = \|a_i - b_j\|^2$. Then

$$F(t) = t(c_{00} + c_{11} - c_{01} - c_{10}) + (c_{01} + c_{10}).$$

Hence an interior solution $t \in (0,1)$ (i.e. all four $x_{ij} > 0$) can occur *only if*

$$c_{00} + c_{11} = c_{01} + c_{10}. \tag{$\star$}$$

Therefore, under any nondegeneracy condition that ensures ($\star$) fails, the minimizer is attained at $t = 0$ or $t = 1$, and at least one of $x_{00}, x_{01}, x_{10}, x_{11}$ is zero.

For the squared Euclidean cost,

$$c_{00} + c_{11} - c_{01} - c_{10} = -2 \langle a_0 - a_1, b_0 - b_1 \rangle,$$

so ($\star$) is equivalent to $\langle a_0 - a_1, b_0 - b_1 \rangle = 0$.

6) No crossings in the planar $p = 1$ case

Assume $d = 2$, $p = 1$, and the segments $[a_1, b_2]$ and $[a_2, b_1]$ intersect at an interior point c and the four points are not collinear.

By the triangle inequality,

$$\|b_1 - a_1\| \leq \|b_1 - c\| + \|c - a_1\|, \qquad \|b_2 - a_2\| \leq \|b_2 - c\| + \|c - a_2\|,$$

with strict inequality in at least one of them because c does not lie on both segments $[a_1, b_1]$ and $[a_2, b_2]$ when the configuration is non-collinear and crossing. Adding gives

$$\|b_1 - a_1\| + \|b_2 - a_2\| < (\|b_1 - c\| + \|c - a_2\|) + (\|b_2 - c\| + \|c - a_1\|) = \|b_1 - a_2\| + \|b_2 - a_1\|,$$

which is the desired inequality.

This strict inequality says that sending $a_1 \to b_1$ and $a_2 \to b_2$ is strictly cheaper than the crossing assignment $a_1 \to b_2$, $a_2 \to b_1$. Therefore an optimal plan cannot allocate positive mass to both crossing pairs simultaneously, hence $x_{10}x_{01} = 0$.