Optimisation, contrôle, données

April 17th, 2026

Responsable: Dario Prandi, dario.prandi@centralesupelec.fr
Chargé de TD: Kevin Le Balc'h, kevin.le-balc-h@inria.fr

Exercise 1: Constrained optimization and Fenchel duality

We consider the optimization problem (P)

$$\min_{(x,y)\in\mathbb{R}^2} x^2 + 2y^2 - 2x - 4y,$$

subject to the constraints

$$x + y \leq 1, \qquad x - y \leq 0.$$

1. Show that (P) admits a unique solution.
2. Write the Lagrangian associated with (P) and state the KKT conditions.
3. Verify that Slater's condition holds. Solve (P) using KKT's theorem.
4. Write the dual problem and verify strong duality.

Solution

1. The objective $f(x, y) = x^2 + 2y^2 - 2x - 4y$ is strictly convex (its Hessian is $\operatorname{diag}(2, 4) \succ 0$), hence continuous and coercive. The constraint set is a closed convex polyhedron and is non-empty (e.g., $(0, 0)$ is feasible). Therefore, (P) admits a unique solution.

2. The Lagrangian is

$$L(x, y, \lambda_1, \lambda_2) = x^2 + 2y^2 - 2x - 4y + \lambda_1(x + y - 1) + \lambda_2(x - y), \quad \lambda_1, \lambda_2 \geq 0.$$

   The KKT conditions are:

   - Primal feasibility: $x + y \leq 1$, $x - y \leq 0$.
   - Dual feasibility: $\lambda_1, \lambda_2 \geq 0$.
   - Complementary slackness: $\lambda_1(x + y - 1) = 0$, $\lambda_2(x - y) = 0$.
   - Stationarity:
$$\begin{cases} 2x - 2 + \lambda_1 + \lambda_2 = 0, \\ 4y - 4 + \lambda_1 - \lambda_2 = 0. \end{cases}$$

3. Slater's condition is satisfied: the point $(0, 1/2)$ satisfies $0 + 1/2 < 1$ and $0 - 1/2 < 0$, so both constraints are strictly satisfied.

   We solve the KKT system by cases.

   Case 1: $\lambda_1 = 0$, $\lambda_2 = 0$. Stationarity gives $x = 1$, $y = 1$. Primal feasibility: $1 + 1 = 2 > 1$. **Infeasible.**

   Case 2: $\lambda_1 > 0$, $\lambda_2 = 0$ (first constraint active, second inactive). From $x+y = 1$ we get $x = 1-y$. Stationarity gives:

$$2(1 - y) - 2 + \lambda_1 = 0 \implies \lambda_1 = 2y, \qquad 4y - 4 + \lambda_1 = 0 \implies 4y - 4 + 2y = 0 \implies y = \frac{2}{3}.$$

   Hence $x = \frac{1}{3}$, $\lambda_1 = \frac{4}{3}$. Primal feasibility: $x - y = \frac{1}{3} - \frac{2}{3} = -\frac{1}{3} \leq 0$ ✓. Dual feasibility: $\lambda_1 = \frac{4}{3} > 0$ ✓.

   Case 3: $\lambda_1 = 0$, $\lambda_2 > 0$ (second constraint active, first inactive). From $x = y$. Stationarity: $2x - 2 + \lambda_2 = 0$ and $4x - 4 - \lambda_2 = 0$. Adding: $6x - 6 = 0 \implies x = 1$, so $y = 1$. Primal feasibility: $x + y = 2 > 1$. **Infeasible.**

Case 4: $\lambda_1 > 0$, $\lambda_2 > 0$ (both constraints active). From $x + y = 1$ and $x = y$: $x = y = \frac{1}{2}$. Stationarity: $\lambda_1 + \lambda_2 = 1$ and $\lambda_1 - \lambda_2 = 2$, giving $\lambda_1 = \frac{3}{2}$, $\lambda_2 = -\frac{1}{2} < 0$. **Dual infeasible.**

The only KKT point is $(x^\star, y^\star) = \left(\frac{1}{3}, \frac{2}{3}\right)$ with multipliers $\lambda_1^\star = \frac{4}{3}$, $\lambda_2^\star = 0$, and optimal value

$$f\left(\tfrac{1}{3}, \tfrac{2}{3}\right) = \tfrac{1}{9} + \tfrac{8}{9} - \tfrac{2}{3} - \tfrac{8}{3} = 1 - \tfrac{10}{3} = -\tfrac{7}{3}.$$

4. The dual function is

$$g(\lambda_1, \lambda_2) = \inf_{(x,y)\in\mathbb{R}^2} L(x, y, \lambda_1, \lambda_2).$$

Since L is strictly convex in (x, y), the infimum is attained at

$$x = \frac{2 - \lambda_1 - \lambda_2}{2}, \qquad y = \frac{4 - \lambda_1 + \lambda_2}{4}.$$

Substituting back:

$$g(\lambda_1, \lambda_2) = -\frac{(2 - \lambda_1 - \lambda_2)^2}{4} - \frac{(4 - \lambda_1 + \lambda_2)^2}{8} - \lambda_1 = -3 + \lambda_1 - \frac{3\lambda_1^2 + 2\lambda_1\lambda_2 + 3\lambda_2^2}{8}.$$

The dual problem is $\max_{\lambda_1,\lambda_2\geq 0} g(\lambda_1, \lambda_2)$. On the face $\lambda_2 = 0$: $g(\lambda_1, 0) = -3 + \lambda_1 - \frac{3}{8}\lambda_1^2$, whose maximum is at $\lambda_1 = \frac{4}{3}$ (from $1 - \frac{3}{4}\lambda_1 = 0$), giving

$$g\left(\tfrac{4}{3}, 0\right) = -3 + \tfrac{4}{3} - \tfrac{3}{8} \cdot \tfrac{16}{9} = -3 + \tfrac{4}{3} - \tfrac{2}{3} = -\tfrac{7}{3}.$$

One checks this is the global maximum (the unconstrained critical point has $\lambda_2 < 0$, and g is decreasing in λ_2 for $\lambda_2 \geq 0$).

Strong duality holds since Slater's condition is satisfied: $f(x^\star, y^\star) = -\frac{7}{3} = g(\lambda_1^\star, \lambda_2^\star)$.

Exercise 2: Projected gradient descent and convergence

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function of class C^1 with L-Lipschitz gradient, and let $C \subset \mathbb{R}^n$ be a non-empty, closed, convex set. We consider the *projected gradient descent* algorithm defined by

$$x^{(k+1)} = \Pi_C\left(x^{(k)} - \alpha\nabla f(x^{(k)})\right),$$

where Π_C denotes the orthogonal projection onto C and $\alpha > 0$ is the step size.

1. Recall why Π_C is well-defined on $\mathbb{R}^n$. Show that, for any $x \in \mathbb{R}^n$ and $y \in C$,

$$\langle x - \Pi_C(x),\, y - \Pi_C(x)\rangle \leq 0.$$

Hint: for the second part, letting $p = \Pi_C(x)$, justify and use the fact that $\|x-p\|^2 \leq \|x-z\|^2$ for $z = p + t(y - p)$, $t \geq 0$.

2. Show that Π_C is 1-Lipschitz, i.e., for all $x, z \in \mathbb{R}^n$,

$$\|\Pi_C(x) - \Pi_C(z)\| \leq \|x - z\|.$$

3. Show that $x^\star \in C$ is a minimizer of f on C if and only if, for any $\alpha > 0$,

$$x^\star = \Pi_C(x^\star - \alpha\nabla f(x^\star)).$$

Hint: reformulate the problem as an unconstrained one, using χ_C, the $0 - \infty$ characteristic function of C. Apply the minimization criterion for the subdifferential and use question 1.

4. Show that, for any $x^\star$ minimizer of f over C and for any $k \geq 0$,

$$\|x^{(k+1)} - x^\star\|^2 \leq \|x^{(k)} - x^\star\|^2 - \alpha\left(\frac{2}{L} - \alpha\right)\|\nabla f(x^{(k)}) - \nabla f(x^\star)\|^2 - 2\alpha\left(f(x^{(k)}) - f(x^\star)\right).$$

Hint: use the characterization of the projection from question 1, and the quadratic upper bound

$$f(y) \leq f(x) + \langle\nabla f(x), y - x\rangle + \frac{L}{2}\|y - x\|^2, \qquad \forall x, y \in \mathbb{R}^n$$

.

5. Show that for $0 < \alpha < 2/L$, the sequence $\{f(x^{(k)})\}$ is non-increasing.
6. Deduce that for $0 < \alpha < 2/L$, the sequence $\{f(x^{(k)})\}$ converges to $f(x^\star)$ and

$$f(x^{(k)}) - f(x^\star) \le \frac{\|x^{(0)} - x^\star\|^2}{2\alpha k}.$$

7. Now consider the specific problem of minimizing $f(x) = \frac{1}{2}\|Ax - b\|^2$ over $C = \{x \in \mathbb{R}^n : x \ge 0\}$ (non-negative least squares), where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$.
   (a) Write ∇f and determine the Lipschitz constant L of the gradient.
   (b) Write explicitly the projected gradient iteration for this problem.

Solution

1. C is a non-empty, closed, convex set. The function $y \mapsto \|y - x\|^2$ is strongly convex and coercive, hence it admits a unique minimizer on C, which defines $\Pi_C(x)$.

   Let $p = \Pi_C(x)$. For any $y \in C$ and $t \in [0, 1]$, $p + t(y - p) \in C$ by convexity. Since p minimizes $\|z - x\|^2$ over C,

$$\|p - x\|^2 \le \|p + t(y - p) - x\|^2 = \|p - x\|^2 + 2t\langle p - x, y - p\rangle + t^2\|y - p\|^2.$$

   Dividing by $t > 0$ and letting $t \to 0$: $\langle p - x, y - p\rangle \ge 0$, i.e., $\langle x - p, y - p\rangle \le 0$.

2. Let $p = \Pi_C(x)$ and $q = \Pi_C(z)$. From question 1:

$$\langle x - p,\, q - p\rangle \le 0 \quad \text{and} \quad \langle z - q,\, p - q\rangle \le 0.$$

   Adding: $\langle x - p - (z - q), q - p\rangle \le 0$, i.e.,

$$\langle (x - z) - (p - q),\, q - p\rangle \le 0 \implies \|p - q\|^2 \le \langle x - z, p - q\rangle \le \|x - z\| \cdot \|p - q\|.$$

   Dividing by $\|p - q\|$ (if $p \ne q$) gives $\|p - q\| \le \|x - z\|$.

3. Minimizing f over C is equivalent to minimizing $F := f + \chi_C$ over $\mathbb{R}^n$. By the theorem on subdifferential characterization of minimizers, $x^\star$ minimizes F iff $0 \in \partial F(x^\star) = \nabla f(x^\star) + \partial\chi_C(x^\star)$.

   By definition of subdifferential applied to χ_C: since $\chi_C(x^\star) = 0$ and $\chi_C(y) = 0$ for all $y \in C$, we have $v \in \partial\chi_C(x^\star)$ iff $\langle v, y - x^\star\rangle \le 0$ for all $y \in C$.

   Hence $x^\star$ is a minimizer iff $-\nabla f(x^\star) \in \partial\chi_C(x^\star)$, i.e.,

$$\langle \nabla f(x^\star),\, y - x^\star\rangle \ge 0, \qquad \forall y \in C.$$

   Setting $z = x^\star - \alpha\nabla f(x^\star)$, this is equivalent to $\langle z - x^\star, y - x^\star\rangle \le 0$ for all $y \in C$, which by question 1 means exactly $x^\star = \Pi_C(z) = \Pi_C(x^\star - \alpha\nabla f(x^\star))$.

4. Let $g_k = \nabla f(x^{(k)})$ and $g^\star = \nabla f(x^\star)$. Since $x^\star$ is a minimizer on C, by question 3,

$$x^\star = \Pi_C(x^\star - \alpha g^\star).$$

   By non-expansiveness of the projection (question 2):

$$\|x^{(k+1)} - x^\star\|^2 \le \|(x^{(k)} - \alpha g_k) - (x^\star - \alpha g^\star)\|^2 = \|x^{(k)} - x^\star\|^2 - 2\alpha\langle g_k - g^\star, x^{(k)} - x^\star\rangle + \alpha^2\|g_k - g^\star\|^2.$$

   For convex L-smooth f, we use

$$f(y) \ge f(x) + \langle \nabla f(x), y - x\rangle + \frac{1}{2L}\|\nabla f(y) - \nabla f(x)\|^2,$$

   with $(x, y) = (x^{(k)}, x^\star)$, which gives

$$\langle g_k, x^{(k)} - x^\star\rangle \ge f(x^{(k)}) - f(x^\star) + \frac{1}{2L}\|g_k - g^\star\|^2.$$

Also, optimality of $x^\star$ on C implies $\langle g^\star, x^{(k)} - x^\star \rangle \geq 0$, hence

$$\langle g_k - g^\star, x^{(k)} - x^\star \rangle \geq f(x^{(k)}) - f(x^\star) + \frac{1}{2L}\|g_k - g^\star\|^2.$$

Substituting into the previous estimate:

$$\|x^{(k+1)} - x^\star\|^2 \leq \|x^{(k)} - x^\star\|^2 - 2\alpha(f(x^{(k)}) - f(x^\star)) - \alpha\left(\frac{2}{L} - \alpha\right)\|g_k - g^\star\|^2.$$

5. For $0 < \alpha < 2/L$, by smoothness,

$$f(x^{(k+1)}) \leq f(x^{(k)}) + \langle g_k, x^{(k+1)} - x^{(k)} \rangle + \frac{L}{2}\|x^{(k+1)} - x^{(k)}\|^2,$$

and by projection optimality with $y = x^{(k)}$,

$$\langle g_k, x^{(k+1)} - x^{(k)} \rangle \leq -\frac{1}{\alpha}\|x^{(k+1)} - x^{(k)}\|^2.$$

Hence

$$f(x^{(k+1)}) \leq f(x^{(k)}) - \left(\frac{1}{\alpha} - \frac{L}{2}\right)\|x^{(k+1)} - x^{(k)}\|^2 \leq f(x^{(k)}).$$

6. For $0 < \alpha < 2/L$, the term $\alpha\left(\frac{2}{L} - \alpha\right) > 0$. Dropping the non-negative gradient term in question 4:

$$2\alpha(f(x^{(k)}) - f(x^\star)) \leq \|x^{(k)} - x^\star\|^2 - \|x^{(k+1)} - x^\star\|^2.$$

Summing from $k = 0$ to $K - 1$ (telescoping):

$$2\alpha \sum_{k=0}^{K-1} (f(x^{(k)}) - f(x^\star)) \leq \|x^{(0)} - x^\star\|^2.$$

Using question 5, $f(x^{(k)})$ is non-increasing, hence

$$f(x^{(K)}) - f(x^\star) \leq \frac{1}{K}\sum_{k=0}^{K-1} (f(x^{(k)}) - f(x^\star)) \leq \frac{\|x^{(0)} - x^\star\|^2}{2\alpha K}.$$

7. (a) $\nabla f(x) = A^\top(Ax - b)$, which is Lipschitz with constant $L = \|A^\top A\| = \sigma_{\max}(A)^2$, the square of the largest singular value of A.

   (b) The projected gradient iteration is

   $$x^{(k+1)} = \max\big(x^{(k)} - \alpha A^\top(Ax^{(k)} - b),\, 0\big),$$

   where $\max(\cdot, 0)$ is applied component-wise.

Exercise 3: Controllability and LQ optimal control

Consider the linear control system in $\mathbb{R}^n$:

$$\dot{x} = Ax + Bu, \qquad x(0) = x_0,$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, and $u \in L^\infty([0,T], \mathbb{R}^m)$. Consider the LQ optimal control problem on $[0,T]$: minimize

$$J(u) = \frac{1}{2}\int_0^T \big(x(t)^\top Q\, x(t) + u(t)^\top R\, u(t)\big)\, dt + \frac{1}{2}x(T)^\top S\, x(T), \qquad \text{(LQ-cost)}$$

where $Q, S \in \mathbb{R}^{n\times n}$ are symmetric positive semi-definite and $R \in \mathbb{R}^{m\times m}$ is symmetric positive definite.

1. Consider the following system:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

   (a) Determine whether the system is controllable, and describe the reachable set from a given initial state x_0.

   (b) State the matrix Riccati equation satisfied by $E(t)$ and give the expression for the optimal control $u^\star(t)$ in feedback form.

   (c) Does the optimal control exist and is it unique, even if the system (A, B) is not controllable? Justify.

2. We now consider the *double integrator* ($n = 2$, $m = 1$):

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad Q = I_2, \qquad R = 1.$$

   (a) Verify that (A, B) is controllable.

   (b) Let $E_T(t)$ denote the solution of the finite-horizon Riccati equation for the cost functional J of Equation (LQ-cost) with $S = 0$. Assume that, for fixed t, $E_T(t)$ converges to a limit matrix $E_\infty \succeq 0$ as $T \to \infty$. Show that E_∞ satisfies the algebraic Riccati equation:

$$Q + A^\top E_\infty + E_\infty A - E_\infty B R^{-1} B^\top E_\infty = 0.$$

   (c) Writing $E_\infty = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, expand the algebraic Riccati equation into a system of scalar equations for a, b, c, d and solve it explicitly.

   (d) Give the optimal feedback control $u^\star = K^\star x$ corresponding to E_∞ and write the closed-loop matrix $A_{\text{cl}} = A - BK^\star$. Verify that A_{cl} is Hurwitz (i.e., is such that the closed-loop system $\dot{x} = A_{\text{cl}} x$ is asymptotically stable).

Solution

1. (a) The Kalman rank condition states that (A, B) is controllable if and only if

$$\operatorname{rank} \mathcal{K}, \qquad \mathcal{K} = \begin{bmatrix} B, \, AB, \, A^2B, \, \dots, \, A^{n-1}B \end{bmatrix} = n.$$

   Here $n = 3$ and we compute:

$$B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad AB = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad A^2B = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

   so $\operatorname{rank} \mathcal{K} = 1 < 3$. The system is **not controllable**.

   The dynamics read $\dot{x}_1 = x_2 + u$, $\dot{x}_2 = x_3$, $\dot{x}_3 = 0$. Thus $x_3(t) = x_3(0)$ is entirely unaffected by u: the reachable set from x_0 is $\{y \in \mathbb{R}^3 : y_3 = x_{0,3}\}$, a plane parallel to the (x_1, x_2)-plane.

   (b) The symmetric positive semi-definite matrix $P(t)$ satisfies the backward Riccati ODE

$$-\dot{E}(t) = Q + A^\top E(t) + E(t)A - E(t)BR^{-1}B^\top E(t), \qquad E(T) = S.$$

   The optimal feedback control is

$$u^\star(t) = -R^{-1}B^\top E(t)\, x(t).$$

   (c) Yes: since $R \succ 0$ the integrand is strictly convex and coercive in u, so J has a unique minimizer for every x_0. Controllability governs whether arbitrary terminal states can be reached; here we are minimizing a cost, not targeting a specific terminal state.

2. (a) The Kalman matrix is

$$\mathcal{K} = [B,\ AB] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

which has rank 2. Hence (A, B) is controllable.

(b) Introduce the time-to-go variable $\tau = T - t$ and define $p(\tau) := P_T(T - \tau)$. Then p solves the autonomous ODE

$$\frac{dp}{d\tau} = Q + A^\top p + pA - pBR^{-1}B^\top p, \qquad p(0) = 0.$$

By assumption, $p(\tau) \to P_\infty$ as $\tau \to +\infty$. For an autonomous ODE, any limit point must be an equilibrium, hence P_∞ satisfies the algebraic Riccati equation.

(c) Observe that P_∞ is symmetric (it is a limit of symmetric matrices), so $c = b$. Write directly

$$P_\infty = \begin{bmatrix} a & b \\ b & d \end{bmatrix}.$$

Then

$$A^\top P_\infty = \begin{bmatrix} 0 & 0 \\ a & b \end{bmatrix}, \quad P_\infty A = \begin{bmatrix} 0 & a \\ 0 & b \end{bmatrix}, \quad P_\infty BR^{-1}B^\top P_\infty = \begin{bmatrix} b^2 & bd \\ bd & d^2 \end{bmatrix}.$$

Hence the ARE gives

$$(1,1):\ 1 - b^2 = 0, \qquad (1,2):\ a - bd = 0, \qquad (2,2):\ 1 + 2b - d^2 = 0.$$

From $(1,1)$, $b = \pm 1$; for the stabilizing positive-semidefinite solution, take $b = 1$. Then $(1,2)$ gives $a = d$, and $(2,2)$ gives $d^2 = 3$, so $d = \sqrt{3}$ (positive branch), hence $a = \sqrt{3}$. Therefore

$$\boxed{P_\infty = \begin{bmatrix} \sqrt{3} & 1 \\ 1 & \sqrt{3} \end{bmatrix}.}$$

This is indeed positive definite: $\det P_\infty = 3 - 1 = 2 > 0$ and $\operatorname{tr} P_\infty = 2\sqrt{3} > 0$.

(d) The optimal feedback law has the form

$$u^\star(x) = -R^{-1}B^\top P_\infty x = -\begin{bmatrix} 1 & \sqrt{3} \end{bmatrix} x.$$

If we denote the (row) gain by

$$K^\star := R^{-1}B^\top P_\infty = \begin{bmatrix} 1 & \sqrt{3} \end{bmatrix},$$

then $u^\star(x) = -K^\star x$ and, with the statement convention $A_{\text{cl}} = A - BK^\star$,

$$A_{\text{cl}} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & \sqrt{3} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -\sqrt{3} \end{bmatrix}.$$

Its characteristic polynomial is

$$\det(\lambda I - A_{\text{cl}}) = \lambda^2 + \sqrt{3}\,\lambda + 1,$$

with roots $\lambda = \frac{-\sqrt{3} \pm i}{2}$. Since $\mathrm{Re}(\lambda) = -\frac{\sqrt{3}}{2} < 0$, the matrix A_{cl} is Hurwitz and the closed-loop system is asymptotically stable.

Exercise 4: Neural ODEs and ResNets

Let $\sigma : \mathbb{R} \to \mathbb{R}$ be a nonlinear activation function applied component-wise. Let $\theta : [0, 1] \to \mathbb{R}^p$ be a time-varying parameter and consider the *neural ODE*

$$\dot{x} = f_\theta(t, x) = W(t)\,\sigma(V(t)\,x + c(t)), \qquad x(0) = x_0 \in \mathbb{R}^n, \tag{1}$$

where $W(t) \in \mathbb{R}^{n \times q}$, $V(t) \in \mathbb{R}^{q \times n}$, $c(t) \in \mathbb{R}^q$, and $\theta(t) = (W(t), V(t), c(t))$ are learnable parameters.

1. Write down the ResNet with L layers corresponding to the Euler discretization of (1) on $[0,1]$ with step $h = 1/L$.
2. Let $\Phi_t : \mathbb{R}^n \to \mathbb{R}^n$ denote the flow map of (1), i.e., $x(t) = \Phi_t(x_0)$. Show that

$$\det\big(D\Phi_t(x_0)\big) > 0, \qquad \forall\, t \in [0,1],\ \forall\, x_0 \in \mathbb{R}^n.$$

   What does this imply about the expressiveness of neural ODEs?
   Hint: use the Liouville–Jacobi formula $\frac{d}{dt}\det(D\Phi_t) = \det(D\Phi_t)\ \mathrm{tr}(D_x f_\theta(t, \Phi_t))$.
3. To overcome this limitation, one considers the *augmented neural ODE*:

$$\frac{d}{dt}\begin{pmatrix} x \\ z \end{pmatrix} = g_\theta(t, x, z), \qquad \begin{pmatrix} x(0) \\ z(0) \end{pmatrix} = \begin{pmatrix} x_0 \\ 0 \end{pmatrix},$$

   Here, $(x, z) \in \mathbb{R}^{n+n'}$ is the augmented state. The function g_θ has the form (1) but with x replaced by (x, z) and learnable parameters of appropriate dimensions. The output map is then $\Psi : x_0 \mapsto x(1)$ (projecting back onto the first n coordinates).

   (a) Explain why Ψ is no longer constrained to have positive Jacobian determinant.
   (b) Provide an explicit example with $n = n' = 1$ and linear activation function σ, where Ψ realizes the map $x_0 \mapsto -x_0$.
   Hint: consider a simple linear system that rotates the state in the (x, z)-plane.

Solution

1. The Euler discretization of (1) with step $h = 1/L$ gives the ResNet

$$x_{k+1} = x_k + \frac{1}{L}\, W_k\, \sigma\big(V_k\, x_k + c_k\big), \qquad k = 0, \ldots, L-1,$$

   where $W_k = W(k/L)$, $V_k = V(k/L)$, and $c_k = c(k/L)$ are the layer weights. This is a residual network: each layer adds a "correction" $\frac{1}{L} W_k \sigma(\cdot)$ to the identity.

2. By the Liouville–Jacobi formula,

$$\frac{d}{dt}\det(D\Phi_t(x_0)) = \det(D\Phi_t(x_0))\ \mathrm{tr}\big(D_x f_\theta(t, \Phi_t(x_0))\big).$$

   Since $\det(D\Phi_0) = \det(\mathrm{Id}) = 1 > 0$, integrating gives

$$\det(D\Phi_t(x_0)) = \exp\left(\int_0^t \mathrm{tr}\big(D_x f_\theta(s, \Phi_s(x_0))\big)\, ds\right) > 0$$

   for all t and x_0, since the exponential is always positive.

   Consequence: the flow map Φ_1 always *preserves orientation*. Therefore, a neural ODE cannot learn any map whose Jacobian has negative determinant anywhere — in particular, it cannot realize reflections, orientation-reversing permutations, or any map of degree -1. This is a fundamental expressiveness limitation of the architecture.

3. (a) The augmented flow $\widetilde{\Phi}_t : (x_0, 0) \mapsto (x(t), z(t))$ in $\mathbb{R}^{n+n'}$ still satisfies $\det(D\widetilde{\Phi}_t) > 0$ by the same argument. However, the map of interest,

$$\Psi = \pi \circ \widetilde{\Phi}_1 \circ \iota,$$

   where $\iota : x_0 \mapsto (x_0, 0)$ (embedding) and $\pi : (x, z) \mapsto x$ (projection), goes between spaces of the *same* dimension n but passes through a higher-dimensional space. Since ι and π are not square maps, the chain rule gives $D\Psi = D\pi \cdot D\widetilde{\Phi}_1 \cdot D\iota$, a product of non-square matrices, whose determinant is unconstrained in sign. In particular, the positivity of $\det(D\widetilde{\Phi}_1)$ does not force $\det(D\Psi) > 0$.

(b) Consider the linear system

$$\frac{d}{dt}\begin{pmatrix} x \\ z \end{pmatrix} = \begin{bmatrix} 0 & -\omega \\ \omega & 0 \end{bmatrix}\begin{pmatrix} x \\ z \end{pmatrix}, \qquad \begin{pmatrix} x(0) \\ z(0) \end{pmatrix} = \begin{pmatrix} x_0 \\ 0 \end{pmatrix},$$

where $\omega > 0$ is a constant. This can be realised as an augmented neural ODE with linear activation $\sigma(s) = s$ and parameters

$$W = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad V = \begin{bmatrix} 0 & -\omega \\ \omega & 0 \end{bmatrix}, \quad c = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The flow of this system is a rotation in the (x, z)-plane:

$$\widetilde{\Phi}_t(x_0, 0) = \begin{bmatrix} \cos(\omega t) & -\sin(\omega t) \\ \sin(\omega t) & \cos(\omega t) \end{bmatrix}\begin{pmatrix} x_0 \\ 0 \end{pmatrix} = \begin{pmatrix} x_0 \cos(\omega t) \\ x_0 \sin(\omega t) \end{pmatrix}.$$

Choosing $\omega = \pi$ gives $\Psi(x_0) = x(1) = -x_0$, realizing the desired map.