Stanford
AA203: Optimal and Learning-based Control
Problem set 1, due on April 15

Problem 1: Consider the following functions:

$$f_1(x, y) = -\log(10 - 2x^2 - y^2)$$
$$f_2(x, y) = x^2(1 + 2y - x^2)$$

(a) Does the origin $(0, 0)$ satisfy the necessary conditions for a local minimum? What about the sufficient conditions?

(b) Is $(0, 0)$ a local minimum? Is it global?

Problem 2: Consider the problem:

$$\begin{array}{rl} \min & x_1 + x_2 \\ \text{subject to} & x_1^2 + x_2^2 = 2. \end{array}$$

(This is the problem of maximizing the perimeter of a rectangle inscribed in a given circle.) By using the Lagrange multiplier theorem, show that it has a unique global maximum and a unique global minimum.

Problem 3: Find all candidates for local minima for the optimization problem:

$$\begin{array}{rl} \min & \frac{1}{2}(x_1^2 + x_2^2 + x_3^2) \\ \text{subject to} & x_1 + x_2 + x_3 \leq -3. \end{array}$$

Problem 4: Consider the shortest path problem in Figure 1, where it is only possible to travel to the right and the numbers represent the travel times for each leg. The control is the decision to go up-right or down-right at each node.

(a) By using Dynamic Programming (DP), find the shortest path from A to B.

(b) Consider a generalized version of the shortest path problem in Figure 1 where the grid has n segments on each side. Find the number of computations required by an exhaustive search algorithm (i.e., the number of routes that such algorithm would need to evaluate) and the number of computations required by a DP algorithm (i.e., the number of DP evaluations). (For example, for the case where $n = 3$, the number of computations for the exhaustive search algorithm is 20 and the number of computations for the DP algorithm is 15.)

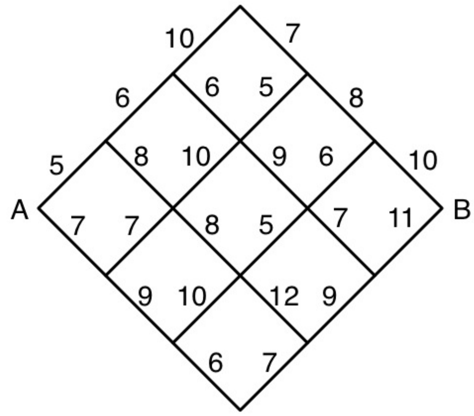

Figure 1: Shortest path problem for Problem 1.

Problem 5: Let $Q \in \mathbb{R}^{n\times n}$ be a symmetric, positive definite matrix, $b \in \mathbb{R}^n$ a given vector and consider the (quadratic) optimization problem

$$\min_{x\in\mathbb{R}^n} f(x) := \frac{1}{2}x^\top Qx - b^\top x. \tag{1}$$

We denote the eigenvalues of Q by $\lambda_1, \dots, \lambda_n$.

(a) Find the unique candidate x^* for a local minimum. Justify why x^* is a global minimum. *Hint: a twice differentiable function f such that $\nabla^2 f(x)$ is positive definite everywhere is strictly convex.*

(b) Show that, starting from any initial point $x^{(0)} \in \mathbb{R}^n$, the Newton's method with $\eta_0 = 1$ converges in one iteration to the optimal solution x^*. Hence, performing one step of the Newton's method is equivalent to solving the linear system of equations $Qx = b$. What would be the downside if n is a large number (say $n \gg 10000$) and the matrix Q has no particular structure?

(c) Let $S \in \mathbb{R}^{n\times n}$ be a symmetric matrix. By the spectral decomposition theorem, there exist an orthogonal matrix $U \in \mathbb{R}^{n\times n}$ and a diagonal matrix $\Sigma = \operatorname{diag}(\mu_1, \dots, \mu_n)$ such that $S = U\Sigma U^\top$. Show that for any $x \in \mathbb{R}^n$, $\|Sx\|_2 = \|\Sigma U^\top x\|_2$. Then, show that for any $z \in \mathbb{R}^n$, $\|\Sigma z\|_2 \le (\max_{i=1,\dots,n} |\mu_i|)\, \|z\|_2$. Conclude that for any $x \in \mathbb{R}^n$, it holds that

$$\|Sx\|_2 \le \left(\max_{i=1,\dots,n} |\mu_i|\right) \|x\|_2.$$

Hint: for an orthogonal matrix $U \in \mathbb{R}^{n\times n}$, it holds that for any $y \in \mathbb{R}^n$, $\|Uy\|_2 = \|U^\top y\|_2 = \|y\|_2$.

(d) Let $\eta > 0$. Show that the eigenvalues of the matrix $(I - \eta Q)$ are equal to $1 - \eta\lambda_1, \dots, 1 - \eta\lambda_n$.

Hint: what is an orthonormal basis of vectors $(v_1, \dots, v_n)$ such that $(I - \eta Q)v_i = (1 - \eta\lambda_i)v_i$ for all $i = 1, \dots, n$?

(e) Consider the gradient method with a constant step size $\eta > 0$, i.e., at each iteration $k \geq 0$, $x^{(k+1)} = x^{(k)} - \eta \nabla f(x^{(k)})$. Define $\delta_k := \|x^{(k)} - x^*\|_2$. Show that for any $k \geq 0$,

$$\delta_{k+1} \leq \gamma(\eta)\delta_k.$$

where $\gamma(\eta) := \max_{i=1,\dots,n} |1 - \eta\lambda_i|$. Show by induction that for any $k \geq 0$,

$$\delta_k \leq \gamma(\eta)^k \delta_0.$$

Deduce a sufficient condition on η so that the sequence $\{x^{(k)}\}$ converges to x^*.

(f) Consider the gradient method with exact line search. At each iteration, denote the descent direction $d_k := -\nabla f(x^{(k)})$ and the optimal step size

$$\eta_k := \operatorname*{argmin}_{\eta \geq 0} f\left(x^{(k)} + \eta d_k\right).$$

Show that

$$\eta_k = \frac{\|d_k\|_2^2}{d_k^\top Q d_k},$$

Let $n = 2$ and $f(x) := \frac{1}{2}(x_1^2 + \gamma x_2^2)$, and $\gamma = 10$.

(g) What is the optimal solution? Implement the gradient method, with constant step size and with exact line search, starting from $x^{(0)} \in \{(5, 1), (1, 5)\}$. What do you observe with exact line search? When does the gradient method "zig zag"? What issue do you observe with constant step size? Repeat the same experiments for $\gamma = 2$. Submit your plots.

Problem 6: Consider the linear time-invariant dynamical system

$$x_{t+1} = Ax_t + Bu_t,$$

where $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times m}$ are given matrices, $x_t \in \mathbb{R}^n$ is the state of the system and $u_t \in \mathbb{R}^m$ the control applied to the system, at time $t \geq 0$.

Let x_0 be a fixed initial state. The goal is to find a sequence of controls $u^* := (u_0^*, \dots, u_{T-1}^*)$ that minimizes the following quadratic cost

$$J(u) := x_T^\top Q_T x_T + \sum_{t=0}^{T-1} x_t^\top Q x_t + u_t^\top R u_t,$$

where $Q, Q_T \in \mathbb{R}^{n \times n}$ and $R \in \mathbb{R}^{m \times m}$ are positive definite matrices. Later in the class, we will see how dynamic programming techniques can be used to derive an elegant solution to the above problem. Before that, we focus on another approach, based on a least-squares formulation or, equivalently, convex quadratic optimization, as studied in Problem 4. Specifically, we reformulate the previous optimization problem into the following form

$$\min_{u \in \mathbb{R}^{mT}} \frac{1}{2} u^\top \widetilde{Q} u - \widetilde{b}^\top u,$$

where $u = (u_0, \ldots, u_{T-1})^\top$, $\widetilde{Q} \in \mathbb{R}^{mT \times mT}$ is a positive definite matrix, and $\widetilde{b} \in \mathbb{R}^{mT}$.

- Write down the matrix $\widetilde{Q}$ and the vector $\widetilde{b}$ in terms of Q, Q_T, R, A, B and x_0.
- By using this reformulation, implement the gradient algorithm of your choice to find the optimal sequence of controls u^* for $Q = I_2$, $Q_T = 10 I_2$, $R = I_1$, $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $x_0 = (1, 0)$ and $T = 20$ (I_n is the identity matrix in dimension n). What is the optimal cost?

Learning goals for this problem set:

Problem 1: To familiarize with the necessary and sufficient conditions for unconstrained optimization.

Problem 2: To familiarize with the Lagrange multiplier theorem for constrained optimization. Also, to appreciate the power of the Lagrange multiplier theorem to derive basic geometrical results.

Problem 3 To familiarize with the KKT conditions for constrained optimization.

Problem 4: To familiarize with the DP algorithm and to appreciate the computational savings of DP versus an exhaustive search algorithm.

Problem 5 To gain insights into the implementation of gradient methods and review some notions of linear algebra.

Problem 6 To familiarize with Linear Quadratic control, and learn a first algorithmic approach to this problem.