Solution guide to the mock Exam

March 2023

This solution guide is provided to cross-check your solutions, detailed explanation is omitted.

Question 1

Part (i): Consider the function $h : \mathbb{R}^2 \to \mathbb{R}$ given by

$$h(x_1, x_2) = \frac{1}{16}x_1^2 + 4x_2^2 + x_1 g(x_2)$$

where the scalar function $g(x_2)$ satisfies

$$\frac{\partial^2 g}{\partial x_2^2}(x_2) = 0, \quad \forall x_2 \in \mathbb{R}.$$

Under what additional condition on g, the function h is strictly convex?

Next, compute the minimum of h for the case $g(x) = \frac{1}{2}(x + 6)$.

$$\frac{\partial h}{\partial x} = \begin{bmatrix} \frac{1}{8}x_1 + g(x_2) \\ 8x_2 + x_1 \frac{\partial g}{x_2} \end{bmatrix}, \qquad \frac{\partial^2 h}{\partial x^2} = \begin{bmatrix} \frac{1}{8} & \frac{\partial g}{x_2} \\ * & 8 \end{bmatrix}.$$

Therefore, h is strictly convex if

$$(\frac{\partial g}{x_2})^2 \leq 1 \Longrightarrow |(\frac{\partial g}{x_2})| \leq 1$$

As for the minimum of h, we compute

$$\frac{1}{8}x_1^* + \frac{1}{2}x_2^* + 3 = 0, \quad 8x_2^* + \frac{1}{2}x_1^* = 0$$

$$\Rightarrow (x_1^*, x_2^*) = (-32, 2)$$

Part (ii): Let A be a full row-rank $m \times n$ matrix with $m < n$, and $b \in \mathbb{R}^n$. Then, the linear equation $Ax = b$ has infinitely many solutions. Among all these solutions, we would like to find the one with minimum norm. This gives rise to the minimization problem

$$\begin{aligned} \text{Minimize}_{x \in \mathbb{R}^n} \ & \frac{1}{2}x^\top x \\ s.t. \quad & Ax = b. \end{aligned}$$

Compute the optimal value $\frac{1}{2}(x^*)^T x^*$by maximizing the *dual problem*. **Hint:** The matrix $AA^\top$ is invertible.

$$\mathcal{L}(x,\lambda) = \frac{1}{2}x^T x + \lambda^T(Ax - b)$$

To compute the dual function, we do

$$\frac{\partial \mathcal{L}}{\partial x} = 0 \to x = -A^T\lambda$$

$$g(\lambda) = \mathcal{L}(x,\lambda)|_{x=-A^T\lambda} = -\frac{1}{2}\lambda^T AA^\top\lambda - \lambda^T b$$

$$-AA^\top\lambda - b = 0 \to \lambda^* = -(AA^\top)^{-1}b.$$

$$\frac{1}{2}(x^*)^T x^* = g(\lambda^*) = \cdots = \frac{1}{2}b^T(AA^T)^{-1}b$$

Part (iii): Consider the function

$$f(x_1,x_2,x_3) = (x_1 - a)^2 + x_2^2 + x_1(x_2 - a) + (x_3 - 2)^2 - \gamma\ln(x_3 + 1)$$

where a is the last digit of your student number, $\gamma \geq 0$, and $\ln(\cdot)$ denotes the natural logarithm. Write a gradient descent algorithm whose solutions converge to the minimizer of f.

$$\begin{aligned}
\dot{x}_1 &= -2(x_1 - a) - (x_2 - a) \\
\dot{x}_2 &= -2x_2 - x_1 \\
\dot{x}_3 &= -2(x_3 - 2) + \frac{1}{x_3 + 1}
\end{aligned}$$

Part (iv): Add the linear constraints

$$\begin{aligned}
x_1 + 2x_2 &= 5 \\
3x_2 + 4x_3 &= 6
\end{aligned}$$

to the problem in Part (iii). Construct a primal-dual algorithm whose solutions converge to the minimizer of f under the above constraints.

$$\mathcal{L}(x,\lambda) = (x_1 - a)^2 + x_2^2 + x_1(x_2 - a) + (x_3 - 2)^2 - \ln(x_3 + 1) + \lambda_1(x_1 + 2x_2 - 5) + \lambda_2(3x_2 + 4x_3 - 6)$$

$$\begin{aligned}
\dot{x}_1 &= -2(x_1 - a) - (x_2 - a) - \lambda_1 \\
\dot{x}_2 &= -2x_2 - x_1 - 2\lambda_1 - 3\lambda_2 \\
\dot{x}_3 &= -2(x_3 - 2) + \frac{1}{x_3 + 1} - 4\lambda_2 \\
\dot{\lambda}_1 &= x_1 + 2x_2 - 5 \\
\dot{\lambda}_2 &= 3x_2 + 4x_3 - 6
\end{aligned}$$

Note: the blue terms show the new terms with respect to part (iii). It saves you time not to recalculate all the derivatives.

Question 2 Let the cost function of a prosumer be given by

$$J(x) = \frac{1}{2}q_1 x^2 + c_1 x$$

and its utility function be given by

$$U(x) = q_2 \ln(x + c_2)$$

where c_1, c_2, q_1, q_2 are all positive constant, and x is a scalar decision variable. The aim is to minimize the *net* cost $F(x) = J(x) - U(x)$. Suppose you have two processors to solve this optimization problem .

Part (i): Formulate this as a distributed optimization problem, where processor 1 uses the cost parameters c_1 and q_1, and processor 2 uses parameters c_2 and q_2.

$$\begin{aligned} \text{Minimize } & \frac{1}{2}q_1 y_1^2 + c_1 y_1 - q_2 \log(y_2 + c_2) \\ s.t. \quad & y_1 = y_2 \end{aligned}$$

Part (ii): By using the Laplacian matrix, write down a primal-dual algorithm to solve the formulated distributed optimization problem.

$$L = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

$$\mathcal{L}(y, \lambda) = \frac{1}{2}q_1 y_1^2 + c_1 y_1 - q_2 \ln(y_2 + c_2) + \begin{bmatrix} \lambda_1 & \lambda_2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$\begin{aligned} \dot{y}_1 &= -q_1 y_1 - c_1 - (\lambda_1 - \lambda_2) \\ \dot{y}_2 &= +\frac{q_2}{y_2 + c_2} + (\lambda_1 - \lambda_2) \\ \dot{\lambda}_1 &= y_1 - y_2 \\ \dot{\lambda}_2 &= -y_1 + y_2 \end{aligned}$$

Part (iii): Set $q_1 = 1$, and $q_2 = 16$. Take both $c_1 = c_2 = c$. Compute algebraically the optimal point x^* minimizing $F(x)$.

$$\frac{\partial F}{\partial x}(x^*) = 0 \rightarrow \frac{16}{x^* + c} = x^* + c \Longrightarrow x^* = 4 - c$$

Question 3 Suppose we modify a resource allocation problem as

$$\begin{aligned} \text{Minimize}_{p_1,p_2,p_3,p_4} \quad & f_1(p_1) + f_2(p_2) + f_3(p_3) + (p_4 - 2)^2 \\ \text{subject to} \quad & \\ & p_1 + p_2 + p_3 + p_4 = d_1 + d_2 + d_3 + d_4 \end{aligned}$$

with $d_1 = 1$, $d_2 = 2$, $d_3 = 3$, $d_4 = 4$, where the blue terms indicate the modifications. How the distributed algorithm for the original problem should be be revised to provide a distributed solution to this modified resource allocation problem with 4 nodes? Note: (i) It is assumed a Laplacian matrix is used in both cases. (ii) You should not compute the whole revised algorithm, and instead should argue which terms should be added to the original algorithm as a result of the addition of the blue terms.

We consider a line graph where node 4 is connected to node 3. Thus, the Laplace matrix of the graph becomes

$$L = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

Therefore, two revisions are needed: Te dynamics related to node 3 should be modified (due to the addition of the link $3 \leftrightarrow 4$) and the dynamics related to node 4 must be added. The updated algorithm is given below where "..." denotes the terms in the original algorithm without the blue terms:

$$\begin{aligned} \dot{p}_1 &= ... \\ \dot{p}_2 &= ... \\ \dot{p}_3 &= ... \\ \dot{p}_4 &= -2(p_4 - 2) + \lambda_4 \\ \dot{\mu}_1 &= ... \\ \dot{\mu}_2 &= ... \\ \dot{\mu}_3 &= ... - \lambda_3 + \lambda_4 \\ \dot{\mu}_4 &= \lambda_3 - \lambda_4 \\ \dot{\lambda}_1 &= ... \\ \dot{\lambda}_2 &= ... \\ \dot{\lambda}_3 &= ... + \mu_3 - \mu_4 \\ \dot{\lambda}_4 &= -\mu_3 + \mu_4 - p_4 + 4 \end{aligned}$$