Grading manual for the Exam of the course: Distributed optimization in engineering systems

April 2020

The response should be explicit, in view of the announcement Update-3 Item 6.

Question 1 (40pt) Consider the function

$$f(x_1, x_2, x_3) = 2x_1^2 + 8x_2^2 + \alpha x_1 x_2 - 18x_1 - 84x_2 + q_3 x_3^2 + x_3$$

for some scalars α and q_3. Choose q_3 as the last digit of your student number. (For example, if your student number is s1234567, then set $q_3 = 7$.)

Part (i): Determine the set of values of α such that the function f is *strictly* convex.

Hessian: (can be replaced by other strict convexity tests)

$$\frac{\partial^2 f}{\partial x^2} = \begin{bmatrix} 4 & \alpha & 0 \\ \alpha & 16 & 0 \\ 0 & 0 & 2q_3 \end{bmatrix}. \qquad ([\mathbf{2pt}])$$

Hessian is positive definite if and only if

$$|\alpha| < 8 \qquad ([\mathbf{2pt}])$$

Part (ii): Determine the set of values of α such that the function f is convex.

$|\alpha| \leq 8$ ([**2pt**])

Part (iii): For the rest of Question 1, set $\alpha = 2$. Compute the global minimum of f.

$$x_1^* = 2, \qquad x_2^* = 5, \qquad x_3^* = -\frac{1}{2q_3} \qquad ([\mathbf{5pt}])$$

If only the equations

$$4x_1 + 2x_2 - 18 = 0, \quad 16x_2 + 2x_1 - 84 = 0, \qquad 2q_3 x_3 + 1 = 0$$

are correct then [3pt=1+1+1].

Part (iv): Write a gradient descent algorithm whose solution converges to the global minimum of f.

$$\dot{x}_1 = -(4x_1 + 2x_2 - 18), \qquad ([\mathbf{2pt}])$$
$$\dot{x}_2 = -(16x_2 + 2x_1 - 84) \qquad ([\mathbf{2pt}])$$
$$\dot{x}_3 = -(2q_3x_3 + 1) \qquad ([\mathbf{2pt}])$$

Part (v): Add the budget constraint

$$2x_1 + 6x_2 = m \tag{1}$$

for some positive number m. (Note that x_3 does not appear in the constraint). Consider the problem of minimizing f under the above constraint.
What should be the amount of budget m in order to have $x_1^* = 3$ in the optimal solution?
Include the optimal values of the Lagrangian multiplier λ^* , the second decision variable x_2^*, and the third decision variable x_3^* in your answer.

$$\mathcal{L}(x,\lambda) = 2x_1^2 + 8x_2^2 + 2x_1x_2 - 18x_1 - 84x_2 + q_3x_3^2 + x_3 + \lambda(2x_1 + 6x_2 - m) \qquad ([\mathbf{2pt}])$$

(If x_3 is not included in $\mathcal{L}$, but the correct argument for not including it is provided full grade for this part)

$$|\lambda| = 3 \qquad ([\mathbf{2pt}])$$
$$x_2^* = 6 \qquad ([\mathbf{2pt}])$$
$$m = 42 \qquad ([\mathbf{3pt}])$$

x_3^* does not change [**2pt**].
If values above are incorrect, then [1pt] for each of the equations

$$2x_2 + 2\lambda = 6$$
$$16x_2 + 6\lambda = 78$$
$$6 + 6x_2 = m$$

(the equations must be only in terms of x_2 and λ in order to get points.)

Note: If you were not able to determine m in part (v), you can write the algorithms in parts (vi) and (vii) in terms of m.

In grading, replace m by the value that the student has obtained in part (v)

Part (vi): Write a primal-dual algorithm to minimize f subject to the budget constraint in in Part (v).

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} -4 & -2 & 0 & -2 \\ -2 & -16 & 0 & -6 \\ 0 & 0 & -2q_3 & 0 \\ 2 & 6 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda \end{bmatrix} + \begin{bmatrix} 18 \\ 84 \\ -1 \\ -m \end{bmatrix}$$

[6pt] *(-2pt for each mistake)*

Part (vii): Write an augmented primal-dual algorithm to minimize f subject to the budget constraint in Part (v).

$$\mathcal{L}(x, \lambda) = 2x_1^2 + 8x_2^2 + 2x_1x_2 - 18x_1 - 84x_2 + q_3x_3^2 + x_3 + \lambda(2x_1 + 6x_2 - m) + \frac{\rho}{2}(2x_1 + 6x_2 - m)^2$$

(**[2pt]**)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} -4-4\rho & -2-12\rho & 0 & -2 \\ -2-12\rho & -16-36\rho & 0 & -6 \\ 0 & 0 & -2q_3 & 0 \\ 2 & 6 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda \end{bmatrix} + \begin{bmatrix} 18-2\rho m \\ 84-6\rho m \\ -1 \\ -m \end{bmatrix}$$

[4pt=2+2] *(for x_1, x_2, respectively, with focus on the terms containing ρ)*
Mistakes in x_3 and λ should be penalized -1 or -2 (each) depending on the mistake.

Question 2 (25 pt) 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

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

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.

$$\text{Minimize } \frac{1}{2} q_1 y_1^2 + c_1 y_1 - q_2 \log(y_2 + c_2) \qquad (\textbf{[3pt]})$$

$$s.t. \quad y_1 = y_2 \qquad (\textbf{[2pt]})$$

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

$$L = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \qquad (\textbf{[1pt]})$$

$$\mathcal{L}(y, \lambda) = \frac{1}{2} q_1 y_1^2 + c_1 y_1 - q_2 \log(y_2 + c_2) \qquad (\textbf{[1pt]})$$

$$+ \begin{bmatrix} \lambda_1 & \lambda_2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \qquad (\textbf{[2pt]})$$

$$\dot{y}_1 = -q_1 y_1 - c_1 - (\lambda_1 - \lambda_2) \qquad (\textbf{[1pt]})$$

$$\dot{y}_2 = +\frac{q_2}{y_2 + c_2} + (\lambda_1 - \lambda_2) \qquad (\textbf{[1pt]})$$

$$\dot{\lambda}_1 = y_1 - y_2 \qquad (\textbf{[1pt]})$$

$$\dot{\lambda}_2 = -y_1 + y_2 \qquad (\textbf{[1pt]})$$

Part (iii): Set $q_1 = 1$, and $q_2 = 16$. Take both c_1 and c_2 equal to the last digit of your student number. Compute algebraically the optimal point x^* minimizing $F(x)$.

$$\frac{16}{x^* + c} = x^* + c \Longrightarrow x^* = 4 - c \qquad (\textbf{[3pt]})$$

Part(iv) Prove (local) convergence of the primal variables of your algorithm to the point computed in the previous part. **Note:** you can solve this part independently of the specific value obtained in part (iii) by carrying out Lyapunov analysis around an equilibrium of your algorithm.

$$V = \frac{1}{2}(y_1 - y_1^*)^2 + \frac{1}{2}(y_1 - y_2^*)^2 + \frac{1}{2}(\lambda_1 - \lambda_1^*)^2 + \frac{1}{2}(\lambda_1 - \lambda_1^*)^2, \qquad ([\mathbf{1pt}])$$

[1pt]→ incremental dynamics. (if not provided, but the analysis is fine, gets the grade)

$$\dot{V} = \cdots = -(y_1 - x^*)^2 + (y_2 - x^*)^2(\frac{16}{y_2 + c} - \frac{16}{x^* + c}) \qquad ([\mathbf{4pt}])$$

[2pt] Convexity or monotonicty argument for showing that the second term is nonpositive.
[1pt] Since $\dot{V} \leq 0$, by applying LaSalle invariance principle, we have $y_1 = y_1^* = x^*$ and $y_2 = y_2^* = x^*$.

Question 3 (20pt) Consider the resource allocation problem

$$\begin{aligned}&\text{Minimize}_{p_1,p_2,p_3} \quad \frac{1}{2}p_1^2 - 2p_1 + (p_2-1)^2 + p_3^4\\ &\text{subject to}\\ &\qquad\qquad p_1+p_2+p_3 = d_1+d_2+d_3\end{aligned}$$

with $d_1 = 1$, $d_2 = 2$, $d_3 = 3$. **Part (i)** By using an incidence matrix, explicitly write a distributed primal-dual algorithm that computes the optimal solution to the above problem.

$$B = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix} \quad \text{or} \quad B = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \quad \text{or with different orientations} \qquad ([\mathbf{2pt}])$$

Reformulating the constraint:

$$\begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix}\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix} = \begin{bmatrix}p_1\\ p_2\\ p_3\end{bmatrix} - \begin{bmatrix}1\\ 2\\ 3\end{bmatrix} \qquad ([\mathbf{2pt}])$$

$$\mathcal{L}(p,\mu,\lambda) = \frac{1}{2}p_1^2 - 2p_1 + (p_2-1)^2 + p_3^4 \qquad ([\mathbf{1pt}])$$

$$+ \begin{bmatrix}\lambda_1 & \lambda_2 & \lambda_3\end{bmatrix}\left(\begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix}\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix} - \begin{bmatrix}p_1\\ p_2\\ p_3\end{bmatrix} + \begin{bmatrix}1\\ 2\\ 3\end{bmatrix}\right). \qquad ([\mathbf{2pt}])$$

Then,

$$\begin{aligned}\dot p_1 &= -p_1 + 2 + \lambda_1 && ([\mathbf{1pt}])\\ \dot p_2 &= -2(p_2-1) + \lambda_2 && ([\mathbf{1pt}])\\ \dot p_3 &= -4p_3^3 + \lambda_3 && ([\mathbf{1pt}])\end{aligned}$$

If only the part containing p_i variables are correct, then [1.5pt] out of 3.

$$\begin{bmatrix}\dot\mu_1\\ \dot\mu_2\end{bmatrix} = -\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}\begin{bmatrix}\lambda_1\\ \lambda_2\\ \lambda_3\end{bmatrix} \qquad ([\mathbf{2pt{=}1{+}1}])$$

If only the formulae $-B^\top\lambda$ is provided [0.5pt] out of 2.

$$\begin{bmatrix}\dot\lambda_1\\ \dot\lambda_2\\ \dot\lambda_3\end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix}\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix} - \begin{bmatrix}p_1\\ p_2\\ p_3\end{bmatrix} + \begin{bmatrix}1\\ 2\\ 3\end{bmatrix} \qquad ([\mathbf{3pt{=}1{+}1{+}1}])$$

If only the formulae $B\mu - p + d$ is provided [1pt] out of 3.
For the choice of B with three links, then μ_3 needs to be introduced as well.

Part (ii): (Information structure) Which variables are defined at the nodes and which variables are defined at the links? What information is needed at each node and what information is required at the link?

[1pt] The variables p_1, p_2, and p_3 are defined on the nodes.
[1pt] The variables λ_1, λ_2, and λ_3 are defined on the nodes.
[1pt] The variables μ_1, μ_2 are defined on the links.
[1pt] Each node variable needs its local nodal variables, as well as the variables of the edge connected to node i.
[1 pt] Each edge variable μ_k needs to have access to the relative information $\lambda_i - \lambda_j$ with $k \sim (i, j)$, which are obtained from the nodes on two ends of the links
(The above can be written explicitly using the choice of the incidence matrix.)