C20 Distributed Systems
Example Paper

Kostas Margellos

kostas.margellos@eng.ox.ac.uk

Hilary Term 2020

Problems

1. (a) Consider the minimization problem

$$\min_{x \in [0,2]} \ x$$
$$\text{subject to } x \geq 1.$$

   Compute the associated dual function by dualizing only the constraint $x \geq 1$.

   (b) For an arbitrary $\epsilon > 0$, consider now the minimization problem

$$\min_{x \in [0,2]} \ x + \epsilon x^2$$
$$\text{subject to } x \geq 1.$$

   As in part (a), compute the associated dual function by dualizing only the constraint $x \geq 1$.

   (c) For which of these cases is the dual function differentiable? Should this have been anticipated?

2. Consider the function

$$F(x_1, x_2) = \max \Big\{ (x_1 - 1)^2 + (x_2 + 1)^2, (x_1 + 1)^2 + (x_2 - 1)^2 \Big\},$$

   where x_1, x_2 are scalars.

   (a) Show that F is strictly convex and its minimum is achieved at $(x_1^\star, x_2^\star) = (0, 0)$.

   (b) Provide the main iterations of the Jacobi algorithm applied to the unconstrained minimization problem

$$\min_{x_1, x_2 \in \mathbb{R}} F(x_1, x_2).$$

(c) Show that if the Jacobi algorithm is initialized at $(x_1(0), x_2(0)) = (1,1)$ then it does not converge to the minimum $(0,0)$, but the generated iterates remain at $(1,1)$.
Which assumption is violated so that the Jacobi algorithm does not converge to $(0,0)$ from any initial condition?

3. Consider the proximal minimization algorithm. For a given step-size $c \in \mathbb{R}$ let

$$\Phi_c(y) = \min_{x \in X} F(x) + \frac{1}{2c}\|x - y\|^2$$

be the mapping that achieves the minimum value in the main step of the algorithm. If F is a convex function with respect to x, show that $\Phi_c(y)$ is a convex function with respect to y.

Hint: Recall that the Euclidean norm is a convex function.

4. Suppose that the sets $X_1, \dots, X_m \subset \mathbb{R}^n$ have a non-empty intersection. Consider then the minimization problem

$$\begin{aligned} \min_{x_1,\dots,x_m,z} \quad & \frac{1}{2}\sum_{i=1}^{m} \|x_i - z\|^2 \\ \text{subject to } \quad & z \in \mathbb{R}^n \\ & x_i \in X_i, \ \text{for all } i = 1, \dots, m. \end{aligned}$$

(a) Applying the Gauss-Seidel algorithm to this problem show that it is equivalent to the following main iterations

$$\begin{aligned} z(k+1) &= \frac{1}{m}\sum_{i=1}^{m} x_i(k) \\ x_i(k+1) &= \Pi_{X_i}[z(k+1)], \ i = 1, \dots, m, \end{aligned}$$

where $\Pi_{X_i}[\cdot]$ denotes the projection of its argument on the set X_i, i.e., $\Pi_{X_i}[z(k+1)] = \arg\min_{x_i \in X_i} \|x_i - z(k+1)\|$.

(b) Show that the Augmented Lagrangian algorithm applied to this problem leads to the same update steps.

(c) Provide a geometric interpretation for this minimization problem.

5. Consider the unconstrained minimization problem

$$\min_{x \in \mathbb{R}^n} \frac{1}{2}\|Ax - b\|_2^2 + \rho\|x\|_1,$$

where $A \in \mathbb{R}^{p\times n}$, $b \in \mathbb{R}^p$ and $\rho > 0$ are given. By $\|\cdot\|_2$ and $\|\cdot\|_1$ we denote the Euclidean and first norm, respectively.

(a) Introducing an auxiliary decision vector $z \in \mathbb{R}^n$ as appropriate, show that the main iterations of the Alternating Direction Method of Multipliers (ADMM) applied to this problem take the form

$$\begin{aligned} x(k+1) &= \arg\min_x \frac{1}{2}\|Ax - b\|_2^2 + \lambda(k)^\top x + \frac{c}{2}\|x - z(k)\|_2^2 \\ z(k+1) &= \arg\min_z \rho\|z\|_1 - \lambda(k)^\top z + \frac{c}{2}\|x(k+1) - z\|_2^2 \\ \lambda(k+1) &= \lambda(k) + c(x(k+1) - z(k+1)). \end{aligned}$$

(b) Show that the first ADMM update step admits the closed form expression

$$x(k+1) = \Big(A^\top A + cI\Big)^{-1}(A^\top b + cz(k) - \lambda(k)),$$

where I is an $n \times n$ identity matrix.

(c) Show that for each element $j = 1, \dots, n$ of $z(k+1)$, the second ADMM update step admits the closed form expression

$$z_j(k+1) = \begin{cases} x_j(k+1) + \frac{1}{c}\lambda_j(k) - \frac{\rho}{c} & \text{if } x_j(k+1) + \frac{1}{c}\lambda_j(k) > \frac{\rho}{c}; \\ 0 & \text{if } |x_j(k+1) + \frac{1}{c}\lambda_j(k)| \le \frac{\rho}{c}; \\ x_j(k+1) + \frac{1}{c}\lambda_j(k) + \frac{\rho}{c} & \text{if } x_j(k+1) + \frac{1}{c}\lambda_j(k) < -\frac{\rho}{c}. \end{cases}$$

Hint: Distinguish between the case where $z_j > 0$ and $z_j < 0$ to perform the minimization in the second ADMM update step.

6. Consider the problem equivalence

$$\begin{aligned} &\min_{x_1,\dots,x_m \in \mathbb{R}} \sum_{i=1}^m f_i(x_i) && \Leftrightarrow && \min_{\substack{x_1,\dots,x_m \in \mathbb{R} \\ z_1,\dots,z_m \in \mathbb{R}}} \sum_{i=1}^m f_i(x_i) \\ &\text{subject to } \sum_{i=1}^m x_i = 0 && && \text{subject to } \sum_{i=1}^m z_i = 0 \\ & && && \qquad\qquad x_i = z_i, \text{ for all } i = 1, \dots, m. \end{aligned}$$

(a) Using the "grouping" in x- and z-variables suggested above, provide the main iterations of the ADMM algorithm applied to this problem.

(b) The projection of a vector $\zeta = (\zeta_1, \dots, \zeta_m)$ on the plane $\sum_{i=1}^m z_i = 0$ is given by

$$\Pi_{\{\sum_{i=1}^m z_i = 0\}}[\zeta] = \zeta - \frac{1}{m}\Big(\begin{bmatrix} 1 & \dots & 1 \end{bmatrix} \cdot \zeta \Big) \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix},$$

where $\begin{bmatrix} 1 & \dots & 1 \end{bmatrix} \cdot \zeta$ is the "dot" product between a $1 \times m$ row-vector and ζ. Use this fact to show that the ADMM iterations determined in part (a) take the form

$$x_i(k+1) = \arg\min_{x_i} f_i(x_i) + \lambda(k) x_i + \frac{c}{2}\|x_i - (x_i(k) - \bar{x}(k))\|^2$$
$$\lambda(k+1) = \lambda(k) + c\bar{x}(k+1),$$

where $\bar{x}(k) = \frac{1}{m}\sum_{i=1}^m x_i(k)$.

7. Consider the minimization problem

$$\min_{x \in \mathbb{R}} \quad \alpha(x+1)^2 + \alpha(x-1)^2$$
$$\text{subject to} \quad x \in [-M, M],$$

where $\alpha > 0$ and $1 < M < \infty$. Treat this program as a two-agent problem with $f_1(x) = \alpha(x+1)^2$, $f_2(x) = \alpha(x-1)^2$, and $X_1 = X_2 = [-M, M]$.

(a) Provide the main iterations of the distributed projected gradient algorithm applied to this problem. You may assume that for all iterations $k = 1, \dots,$ the "mixing" weights are given by $a_j^i(k) = \frac{1}{2}$, for all $i, j = 1, 2$.

(b) If the algorithm is initialized at $x_1(0) = -1$ and $x_2(0) = 1$, provide a closed form expression for the updates $x_i(k)$, $i = 1, 2$.

(c) Compare these updates with the ones of the distributed proximal minimization algorithm derived in Lecture 4 of your notes. Which of the two algorithms converges faster to the minimizer $x^\star = 0$?