EE364b Spring 2020 Homework 6

Due Friday 5/14 at 11:59pm via Gradescope

6.1 (4 points) Consider the problem

$$\begin{array}{ll} \text{minimize} & (x_1 - b_1)^2 + (x_2 - b_2)^2 \\ \text{subject to} & x_1 = x_2, \end{array}$$

where $x_1, x_2, b_1, b_2 \in \mathbf{R}$ are scalars. Here x_1 and x_2 are local variables, which need to satisfy the consensus constraint $x_1 = x_2$.

(a) (1 point) Derive the dual decomposition updates for x_1, x_2 and λ where λ is the dual variable that corresponds to the constraint $x_1 = x_2$. (See page 10 of dual decomposition lecture slides)

(b) (1 point) Find the value of the optimal dual parameter λ^* as a function of b_1 and b_2.

(c) (1 point) Show that the dual decomposition method yields dual iterates $\lambda^{(k)}$ that obey

$$\lambda^{(k+1)} - \lambda^* = (1 - \alpha)(\lambda^{(k)} - \lambda^*)$$

where α is the fixed step size in the dual subgradient method update, and k is the iteration counter.

(d) (1 point) Show that the iterates $x_1^{(k)}, x_2^{(k)}, \lambda^{(k)}$ converge to their optimal values for a small enough step size α.

6.2 *(4 points) Distributed ridge regression.* Consider the ℓ_2-regularized least-squares ('ridge regression') problem

$$\text{minimize} \quad f(z) = (1/2) \left\| \begin{bmatrix} A_1 & 0 & B_1 \\ 0 & A_2 & B_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ y \end{bmatrix} - \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \right\|_2^2 + \lambda \left\| \begin{bmatrix} x_1 \\ x_2 \\ y \end{bmatrix} \right\|_2^2,$$

with optimization variable $z = (x_1, x_2, y) \in \mathbf{R}^{n_1} \times \mathbf{R}^{n_2} \times \mathbf{R}^p$. We can think of x_i as the local variable for system i, for $i = 1, 2$; y is the common or coupling variable.

(a) *(2 points) Primal decomposition.* Explain how to solve this problem using primal decomposition, using the subgradient method for the master problem.

(b) *(2 points) Dual decomposition.* Explain how to solve this problem using dual decomposition, using the subgradient method for the master problem. Give a condition (on the problem data) that allows you to guarantee that the primal variables $x_i^{(k)}$ converge to optimal values.

6.3 (8 points) As a part of this homework, you will submit short reviews for the submitted midterm progress reports. On Canvas, we will assign you 2-3 midterm progress reports to be reviewed. You can check the papers assigned to you through the "Midterm Progress Report" assignment on Canvas, which will be available after all the reports submitted. Please read the report all the way through. Summarize the problem and the approach. Point out the strengths and weaknesses. Ask questions you would like addressed in the final report (e.g., comparison with another optimization approach). Please try to give constructive feedback to improve the project results, and suggest alternatives, e.g., different optimization methods, application areas, ideas/techniques to try out.

6.4 *(extra credit, 4 points) Distributed method for bi-commodity network flow problem.* We consider a network (directed graph) with n arcs and p nodes, described by the incidence matrix $A \in \mathbf{R}^{p\times n}$, where

$$A_{ij} = \begin{cases} 1, & \text{if arc } j \text{ enters node } i \\ -1, & \text{if arc } j \text{ leaves node } i \\ 0, & \text{otherwise.} \end{cases}$$

Two commodities flow in the network. Commodity 1 has source vector $s \in \mathbf{R}^p$, and commodity 2 has source vector $t \in \mathbf{R}^p$, which satisfy $\mathbf{1}^T s = \mathbf{1}^T t = 0$. The flow of commodity 1 on arc i is denoted x_i, and the flow of commodity 2 on arc i is denoted y_i. Each of the flows must satisfy flow conservation, which can be expressed as $Ax + s = 0$ (for commodity 1), and $Ay + t = 0$ (for commodity 2).

Arc i has associated flow cost $\phi_i(x_i, y_i)$, where $\phi_i : \mathbf{R}^2 \to \mathbf{R}$ is convex. (We can impose constraints such as nonnegativity of the flows by restricting the domain of ϕ_i to $\mathbf{R}^2_+$.) One natural form for ϕ_i is a function only the total traffic on the arc, *i.e.*, $\phi(x_i, y_i) = f_i(x_i + y_i)$, where $f_i : \mathbf{R} \to \mathbf{R}$ is convex. In this form, however, ϕ is not strictly convex, which will complicate things. To avoid these complications, we will assume that ϕ_i is strictly convex.

The problem of choosing the minimum cost flows that satisfy flow conservation can be expressed as

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^n \phi_i(x_i, y_i) \\ \text{subject to} & Ax + s = 0, \quad Ay + t = 0, \end{array}$$

with variables $x,\ y \in \mathbf{R}^n$. This is the *bi-commodity network flow problem.*

(a) (2 points) Propose a distributed solution to the bi-commodity flow problem using dual decomposition. Your solution can refer to the conjugate functions ϕ_i^*.

(b) (2 points) Use your algorithm to solve the particular problem instance with

$$\phi_i(x_i, y_i) = (x_i + y_i)^2 + \epsilon(x_i^2 + y_i^2), \qquad \mathbf{dom}\, \phi_i = \mathbf{R}^2_+,$$

with $\epsilon = 0.1$. The other data for this problem can be found in `bicommodity_data.[m|jl]`. To check that your method works, compute the optimal value $p^\star$, using CVX.

For the subgradient updates use a constant stepsize of 0.1. Run the algorithm for 200 iterations and plot the dual lower bound versus iteration. With a logarithmic vertical axis, plot the norms of the residuals for each of the two flow conservation equations, versus iteration number, on the same plot.

Hint. We have posted a function `[x,y] = quad2_min(eps,alpha,beta)`, which computes

$$(x^*, y^*) = \underset{x \geq 0, y \geq 0}{\operatorname{argmin}} \left((x+y)^2 + \epsilon(x^2 + y^2) + \alpha x + \beta y \right)$$

analytically. You might find this function useful.