due May 17, Friday 23:59pm

EE364b Homework 6

5.0 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.

(a) Derive the dual decomposition updates for x_1, x_2 and λ where λ is the Lagrange multiplier.

(b) Find the value of the optimal Lagrange multiplier λ^* as a function of b_1 and b_2.

(c) Show that the Lagrange multiplier obeys the linear dynamical system

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

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

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

5.2 *Distributed lasso.* Consider the ℓ_1-regularized least-squares ('lasso') problem

$$\text{minimize} \quad f(z) = (1/2) \left\| \begin{bmatrix} A_1 & 0 & B_1 \\ 0 & A_2 & B_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ y \end{bmatrix} - \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \right\|_2^2 + \lambda \left\| \begin{bmatrix} x_1 \\ x_2 \\ y \end{bmatrix} \right\|_1,$$

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

(b) *Dual decomposition.* Explain how to solve this problem using dual decomposition, using (say) 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.

(c) *Numerical example.* Generate some numerical data as explained below, and solve the problem (using CVX/CVXPY) to find the optimal value $p^\star$. Implement primal and dual decomposition (as in parts (a) and (b)), using CVX/CVXPY to solve the subproblems, and the subgradient method for the master problem in both cases. For primal decomposition, plot the relative suboptimality $(f(z^{(k)}) - p^\star)/p^\star$ versus iteration. For dual decomposition, plot the relative consistency residual $\|y_1^{(k)} - y_2^{(k)}\|_2/\|y^\star\|_2$ versus iteration, where $y^\star$ is an optimal value of y for the problem. In each case, you needn't worry about attaining a relative accuracy better than 0.001, which corresponds to 0.1%.

Generating the data. Generate A_i, B_i, and c_i with entries from a standard Gaussian, with dimensions $n_1 = 100$, $n_2 = 200$, $p = 10$, $m_1 = 250$, and $m_2 = 300$ (these last two are the dimensions of c_1 and c_2). Check that the condition you gave in part (b) is satisfied. Choose $\lambda = 0.1\lambda_{\max}$, where $\lambda_{\max}$ is the value of λ above which the solution is $z^\star = 0$:

$$\lambda_{\max} = \|(A_1^T c_1, A_2^T c_2, B_1^T c_1 + B_2^T c_2)\|_\infty$$

To get reasonable convergence (say, in a few tens of iterations), you may need to play with the subgradient step size.

You are of course welcome (even, encouraged) to also try your distributed lasso solver on problem instances other than the one generated above.