Math Programming II - Homework 4

Problem 1

Let $f : R^n \mapsto R$ be a continuously differentiable function, let X be a closed convex set, and let c be a positive scalar. Show that the proximal algorithm

$$x^{k+1} = \underset{x \in X}{\text{argmin}} \left\{ f(x) + \frac{1}{2c}\|x - x^k\|^2 \right\} \tag{1}$$

is a special case of the block coordinate descent method applied to the problem

$$\begin{aligned} &\text{minimize} \;\; f(x) + \frac{1}{2c}\|x - y\|^2 && (2) \\ &\text{subject to} \quad x \in X,\; y \in R^n, && (3) \end{aligned}$$

which is equivalent to the problem of minimizing f over X.

Problem 2

This exercise shows that for each problem for which there are nonzero Lagrange multipliers, there are infinitely many corresponding problems with no Lagrange multipliers. Consider the problem $\min_{h(x)=0} \; f(x)$, and suppose that x^* is a local minimum such that $\nabla f(x^*) \neq 0$. Show that x^* is a local minimum of the equality constrained problem

$$\min_{\|h(x)\|^\rho = 0} \; f(x),$$

where $\rho > 1$, and that for this problem there are no Lagrange multipliers.

Problem 3

Let x^* be a feasible point that is regular and together with some λ^* satisfies the first- and second-order necessary conditions for the problem

$$\begin{aligned} &\text{minimize} \quad f(x) && (4) \\ &\text{subject to} \quad h_i(x) = 0, \quad i = 1, ..., m && (5) \end{aligned}$$

Show that x^* and λ^* satisfy the second-order sufficient conditions if and only if the matrix

$$\begin{bmatrix} \nabla^2_{xx} L(x^*, \lambda^*) & \nabla h(x^*) \\ \nabla h(x^*)^T & 0 \end{bmatrix}$$

is nonsingular.

Problem 4

Given three numbers n, m, r and a constant matrix $Z \in \mathbb{R}^{n \times m}$, consider the optimization problem

$$\begin{aligned} \underset{X \in \mathbb{R}^{n \times r}, Y \in \mathbb{R}^{r \times m}}{\text{minimize}} \quad & \|Z - XY\|_F^2 \\ \text{subject to} \quad & X \geq 0, \quad Y \geq 0 \end{aligned} \tag{6}$$

(note that the sign "$\geq$" means that all elements of the corresponding matrix are nonnegative, and that $\|\cdot\|_F$ denotes the Frobenius norm).

- Write the first-order optimality conditions for (6).
- Describe how to solve (6) using the gradient projection method with the step size along the feasible direction chosen to be 1 and the step size along the projection arc to be designed using the Armijo rule. Derive the iterative equations and simplify them as much as possible.
- Describe how to solve (6) using the block coordinate descent method where we solve for X and Y alternatively. Discuss (without having to write the equations in details) how to find a global solution (up to a given precision) of each of the optimization sub-problems that needs to solved in the iterations of the block coordinate descent.