Math Programming II - Midterm Exam

Problem 1 (10 points)

Consider an $n \times n$ symmetric matrix Q that is sign indefinite (meaning that it is neither positive semidefinite nor negative semidefinite). Let $x^T Q x$ be minimized using the gradient algorithm with the initial point x^0 and a sufficiently small constant step size t. Find all values x^0 that make the algorithm converge to a saddle point of the function $x^T Q x$.

Problem 2 (25 points)

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

$$\begin{aligned} \underset{X \in \mathbb{R}^{n \times r}, Y \in \mathbb{R}^{r \times m}}{\text{minimize}} \quad & \|Z - XY\|_F^2 \\ \text{subject to} \quad & X \geq 0, \quad Y \geq 0 \end{aligned} \tag{1}$$

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

- (10 points) Write the first-order optimality conditions for (1).
- (10 points) Describe how to solve (1) 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.
- (5 points) Describe how to solve (1) 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 need to solved in the iterations of the block coordinate descent.

Problem 3 (20 points)

Given an $n \times n$ symmetric matrix Q, consider the problem of minimizing $x^T Q x$. Assume that there are three positive numbers α, β, γ such that the eigenvalues of Q all belong to the two intervals $[\alpha, \beta]$ and $[\alpha + \gamma, \beta + \gamma]$. Let $x^0, x^1, x^2, ...$ be a sequence of points generated in such a way that, for every $k \in \{0, 1, 2, ...\}$, x^{k+1} is obtained from x^k using two iterations of the conjugate gradient method (the resulting algorithm can be regarded as conjugate gradient but with a restart every two steps). Show that this algorithm (more precisely, the sequence of cost functions generated by the algorithm) converges linearly with the rate $\left(\frac{\beta-\alpha}{\beta+\alpha}\right)^2$ (hint: you may want to define a polynomial based on the midpoints of the two intervals to bound the improvement due to any two iterations of the conjugate gradient method).

Problem 4 (15 points)

Consider two $n \times n$ symmetric matrices P and Q for which the set $\{(x^T Px, x^T Qx) | x \in \mathbb{R}^n\}$ is convex and of dimension 2. Consider the optimization problem:

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & x^T Q x \\ \text{subject to} \quad & x^T P x = a \end{aligned} \tag{2}$$

where a is a constant. Assume that the problem has a globally optimal solution. Analyze the duality gap of the above optimization problem and find a range of possible values for this gap (feel free to explain your ideas geometrically and with visualization).

Problem 5 (30 points)

Consider the optimization problem:

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & x' M_0 x \\ \text{subject to} \quad & x' M_i x = 1, \quad i = 1, ..., m \end{aligned} \tag{3}$$

where M_0 is non-singular. Let x^* denote a nonzero local minimum of the above optimization problem that satisfies the second-order sufficient condition. To convert the above problem to an unconstrained optimization problem, consider the following optimization:

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad x' M_0 x + c \sum_{i=1}^{m} p(x' M_i x - 1) \tag{4}$$

where $p(\cdot) : \mathbf{R} \to \mathbf{R}$ is a continuous function with the properties that it is always nonnegative and is zero only at the origin, and where c is a positive constant.

- (15 points) Assume that $p(\cdot)$ is differentiable at the origin. Find all penalty terms $p(\cdot)$ for which there exists a constant $\bar{c}$ such that x^* is a local minimum of (4) for all $c > \bar{c}$ (hint: compare the optimality conditions of the constrained and unconstrained optimization problems).
- (15 points) Assume that $p(\cdot)$ is non-differentiable at the origin and let α and β denote its right and left derivatives at the origin. Is there a threshold $\bar{c}$ such that x^* is a local minimum of (4) for all $c > \bar{c}$? If so, find the smallest threshold you can (note: you can't use the exact penalty theorem from the book and need to redo its proof if you decide to use this theorem. Also, feel free to make any assumptions you need).