Nonlinear Optimization (18799 B,PP)
IST-CMU PhD course
Spring 2011
Instructor: `jxavier@isr.ist.utl.pt`
TA: `augustos@andrew.cmu.edu`

Important: The homework is due April 16th

Homework 6

Problem A. *(Convex regularizers)* Consider the optimization problem

$$\text{minimize} \quad \tfrac{1}{2}\left\|Ax - z\right\|^2 + \sigma_C(x) \tag{1}$$

where $C \subset \mathbb{R}^n$ is a closed convex set, $A \in \mathbb{R}^{m\times n}$ and $z \in \mathbb{R}^m$. Recall that σ_C denotes the support function of C. The set C and both A, z are given. The variable to optimize in problem (1) is $x \in \mathbb{R}^n$. Assume that C is compact and that it contains the origin (this implies $\sigma_C(x)$ is finite-valued and nonnegative for all x) and $\operatorname{rank}(A) = m$ (necessarily, $m \le n$).

A possible interpretation of problem (1) is as follows. We want to estimate an information signal $x \in \mathbb{R}^n$ which was distorted by a linear system $A \in \mathbb{R}^{m\times n}$, based on the available noisy system output $z \in \mathbb{R}^m$ (i.e., $z = Ax +$ "noise") and some *a priori* knowledge about x (e.g., x is sparse). The prior knowledge is taken into account by the the regularizer σ_C: it gives preference to some x's over others. For example, as mentioned in class, if it is known that x is sparse then it is appropriate to use a regularizer like $\sigma_C(x) = \alpha\left\|x\right\|_1$ for some $\alpha > 0$ (which corresponds to $C = B_\infty(0, \alpha)$).

(a) Let $p^\star$ be the optimal value of (1). Show that $p^\star = d^\star$ where $d^\star$ is the optimal value of

$$\begin{array}{ll} \text{maximize} & \tfrac{1}{2}\left\|z\right\|^2 - \tfrac{1}{2}\left\|\lambda - z\right\|^2 \\ \text{subject to} & A^\top\lambda \in C \end{array} \tag{2}$$

with optimization variable $\lambda \in \mathbb{R}^m$.
Hint: problem (1) can be reformulated as

$$\begin{array}{ll} \text{minimize} & \tfrac{1}{2}\left\|r\right\|^2 + \sigma_C(x) \\ \text{subject to} & r = Ax - z \end{array} \tag{3}$$

with optimization variable $(x, r) \in \mathbb{R}^n \times \mathbb{R}^m$. Obviously, problems (1) and (3) have the same optimal value $p^\star$. Now, dualize (3).

(b) Note that (2) corresponds to projecting z onto the set

$$S := \left(A^\top\right)^{-1}(C) = \{\lambda \,:\, A^\top\lambda \in C\}.$$

Show that S is non-empty, convex and compact (thus, closed).

(c) Consider the special case $A = I_n$ (usually, (1) is then called a denoising problem). Show that $x^\star = z - p_C(z)$ solves (1), where $p_C(z)$ denotes the projection of z onto C.

Problem B. *(Economic games)* Let a directed graph with n nodes (labeled $i = 1, \ldots, n$) and m directed arcs (labeled $k = 1, \ldots, m$) be given. The first and last nodes are termed the source and the sink, respectively. The kth arc ($k = 1, \ldots, m$) is an ordered pair (i_k, j_k) meaning that the kth arc starts at node i_k and ends at node j_k. The graph is represented by its node-arc incidence matrix $A \in \mathbb{R}^{n \times m}$:

$$A_{ik} = \begin{cases} 1 & \text{, if arc } k \text{ starts at node } i \\ -1 & \text{, if arc } k \text{ ends at node } i \\ 0 & \text{, otherwise.} \end{cases}$$

As an example, the node-arc incidence matrix of the directed graph in figure 1 is

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 \end{bmatrix}.$$

Assume that there are no self-loops (no arc of the form (i, i)) and that there is, at most, one

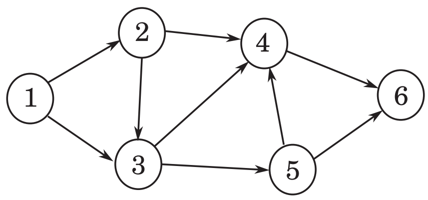

Figure 1: A directed graph with 6 nodes and 9 directed arcs

arc from i to j (no repeated arcs). Additionally, assume that there is at least one directed path from the source to the sink.

Think of the given graph as representing a system of directed pipelines which can carry a given liquid from the source to the sink. Assume that the nodes are flow-preserving: the amount of the liquid which enters a node leaves it. We will be interested in transferring one unit of the liquid from the source to the sink. In general, there are many possible flows through the graph guaranteeing that. Mathematically, a valid flow is a nonnegative vector $x \in \mathbb{R}^m$ where $x_k \geq 0$ ($k = 1, \ldots, m$) denotes the amount of liquid flowing along the direction of arc k. It must obey the balance equations

$$Ax = b, \qquad b := \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{bmatrix}. \tag{4}$$

The first and last linear constraint in (4) mean that one unit of flow is sent by the source and received by the sink, while the intermediate constraints ensure flow is preserved at each node.

Suppose that prices are assigned to each arc, i.e, we are given a nonnegative vector $\alpha \in \mathbb{R}^m$ where α_k ($k = 1, \ldots, m$) denotes the price, say, in dollars, per unit of liquid carried by the pipeline k. Note that, if we are given a price vector $\alpha \in \mathbb{R}^m$, the problem of finding the minimum cost flow corresponds to the optimization problem

$$\begin{array}{ll} \text{minimize} & \alpha^\top x \\ \text{subject to} & Ax = b \\ & x \geq 0. \end{array} \tag{5}$$

In the following, we let $p^\star(\alpha)$ be the optimal value of the optimization problem (5).

We will consider two optimization problems within this framework.

- Suppose we are a customer of the given infrastructure of directed pipelines. We want to use it to transfer one unit of the liquid from source to sink. The rules are as follows: first, we specify a flow x; then, the company owning the infrastructure (and that will carry out the implementation of x) chooses a price vector α and withdraws $\alpha^\top x$ dollars from our bank account. The price vector α chosen by the company is taken from a compact polyhedron $\mathcal{A} = \{\alpha \,:\, F\alpha \leq g\}$. The polyhedron of prices $\mathcal{A}$ is known to us beforehand, i.e., when we choose x (of course, we do not know which point $\alpha \in \mathcal{A}$ is going to be selected after). Naturally, we design the flow x by solving

$$\begin{array}{ll} \text{minimize} & \sup\{\alpha^\top x \,:\, \alpha \in \mathcal{A}\} \\ \text{subject to} & Ax = b \\ & x \geq 0. \end{array} \tag{6}$$

  The interpretation of (6) is clear: we want to find a flow x which minimizes the worst-case cost that we will be charged to us by the company. Let $p^\star$ be the optimal value of (6): $p^\star$ is the minimum amount of dollars we must have in our bank account when we select a flow in order to face any subsequent choice of prices by the owning company.

- Now, suppose we work in the owning company. The rules are now as follows: first, we choose a price vector α and make it public; then, a costumer (knowing α) requests a flow x (which we implement) and we withdraw $\alpha^\top x$ dollars from his bank account. Assume the costumer is smart, i.e., he chooses the minimum cost flow for our price α by solving (5). Thus, we will transfer to our account $p^\star(\alpha)$ dollars (recall that $p^\star(\alpha)$ denotes the optimal value of (5)). In order to maximize our profit, we will naturally set the price vector α by solving

$$\begin{array}{ll} \text{maximize} & p^\star(\alpha) \\ \text{subject to} & \alpha \in \mathcal{A}. \end{array} \tag{7}$$

  Let $d^\star$ be the optimal value of (7): $d^\star$ is the maximum amount of dollars we can extract from a smart costumer.

Show that $p^\star = d^\star$.