EE 25088 Convex Optimization II

Problem Set IV

Due Date: Khordad 24, 1401

1) ***Processor Scheduling with Minimum Energy:*** A single processor can adjust its speed in each of T time periods, labeled as $1, \cdots, T$. Its speed in period t is denoted by s_t, $t = 1, \cdots, T$. The speeds must lie between given minimum and maximum values, $S^{\min}$ and $S^{\max}$, respectively, and must satisfy a slew-rate limit, i.e., $|s_{t+1} - s_t| \leq R$, for $t = 1, \cdots, T-1$. (R is the maximum allowed period-to-period change in speed.) The energy consumed by the processor in period t is given by $\phi(s_t)$, where $\phi : \mathbf{R} \to \mathbf{R}$ is increasing and convex. The total energy consumed over all the periods can be calculated by $E = \sum_{t=1}^{T} \phi(s_t)$.

   The processor must handle n jobs, labeled $1, \cdots, n$. Each job has an availability time $A_i \in \{1, \cdots, T\}$, and a deadline $D_i \in \{1, \cdots, T\}$, with $D_i \geq A_i$. The processor cannot start work on job i until period $t = A_i$, and must complete the job by the end of period D_i. Job i involves a (nonnegative) total workload W_i. You can assume that in each time period, there is at least one job available, i.e., for each t, there is at least one i with $A_i \leq t$ and $D_i \geq t$.

   In period t, the processor allocates its effort across the n jobs as $\boldsymbol{\theta}_t$, where $\mathbf{1}^T \boldsymbol{\theta}_t = 1, \boldsymbol{\theta}_t \succeq \mathbf{0}$. Here $\theta_{t,i}$ (the ith component of $\boldsymbol{\theta}_t$) gives the fraction of the processor effort devoted to job i in period t. Respecting the availability and deadline constraints requires that $\theta_{t,i} = 0$ for $t < A_i$ or $t > D_i$. To complete the jobs we must have $\sum_{t=A_i}^{D_i} \theta_{t,i} s_t \geq W_i$, for $i = 1, \cdots, n$.

   a) Formulate the problem of choosing the speeds $s_1, \cdots, s_T$, and the allocations $\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_T$, in order to minimize the total energy E, as a convex optimization problem. The problem data are $S^{\min}$, $S^{\max}$, R, ϕ, and the job data, A_i, D_i, W_i, $i = 1, \cdots, n$. Be sure to justify any change of variables, or introduction of new variables, that you use in your formulation.

   b) Carry out your method on the problem instance described in `proc_sched_data.m`, with quadratic energy function defined as $\phi(s_t) = \alpha + \beta s_t + \gamma s_t^2$. The required parameters of α, β, and γ are given in the data file.[1]

[1] Executing this file will also demonstrate a plot showing the availability times and deadlines for the jobs.

Give the energy obtained by your speed profile and allocations. Also, plot these using the command `bar((s*ones(1,n)).*Theta,1,'stacked')`, where s is the $T \times 1$ vector of speeds, and Θ is the $T \times n$ matrix of allocations with components $\theta_{t,i}$. This will show, at each time period, how much effective speed is allocated to each job, while the top of the plot demonstrating the speed s_t.

2) Exercise 2.7 a,b,e, and f of [1].

[1] M. Neely, Stochastic Network Optimization with Application to Communication and Queueing Systems. Morgan & Claypool Publishers 2010.

3) Problem 2.2(a) of [2].

[2] A. Shapiro, D. Dentcheva, and A. Ruszczynski, *Lectures on Stochastic Programming: Modeling and Theory.* Society for Industrial and Applied Mathematics (SIAM), 2nd Ed., 2014.

4) Consider the following cost minimization linear program (LP).

$$\underset{x \in \mathbb{R}^n}{\text{minimize}}\ c^T x$$
$$\text{subject to } Ax \preceq b.$$

Here, the cost vector $c \in \mathbb{R}^n$ is random, normally distributed with mean $\mathbf{E}[c] = c_0$ and covariance matrix Σ, defined as $\Sigma = \mathbf{E}[(c - c_0)(c - c_0)^T]$. $A, b,$ and x are, however, deterministic. In the following, we explore some approaches to minimize the aforementioned cost function.

a) Formulate an optimization problem that minimizes the expected cost and explain whether it is an LP problem.

b) In general, there is a tradeoff between small *expected cost* and small *cost variance*. One way to take variance into account is to minimize a linear combination of the expected value $\mathbf{E}\left[c^T x\right]$ and the variance $\mathbf{var}[c^T x] = \mathbf{E}[(c^T x)^2] - \left(\mathbf{E}[c^T x]\right)^2$ as $\mathbf{E}\left[c^T x\right] + \lambda \mathbf{var}[c^T x]$. This is called the 'risk-sensitive cost', and the parameter $\lambda \geq 0$ is called the risk-aversion parameter, since it sets the relative values of cost variance and expected value. For $\lambda > 0$, we are willing to tradeoff an increase in expected cost for a decrease in cost variance. How can one minimize the risk-sensitive cost? Explicitly explain whether this is a convex optimization problem or not by properly formulating the mentioned problem. What can we say for the case of $\lambda < 0$?

c) Another way to deal with the randomness in the cost function $c^T x$ is to formulate the problem as

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n,\beta}{\text{minimize}} & \beta \\ \text{subject to} & \mathbf{prob}\left(c^T x \geq \beta\right) \leq \alpha \\ & Ax \preceq b. \end{array}$$

Here, α is a fixed parameter, which corresponds roughly to the reliability we require, and might typically have a value of 0.01. Is this problem a convex optimization problem? Be as specific as you can.

5) In this problem, we get familiar with quadratic penalty function.

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \cdots, m, \end{array} \quad (1)$$

where the functions $f_i : \mathbf{R}^n \to \mathbf{R}$ are differentiable and convex. The quadratically-penalized function $\phi(x)$ is defined as $\phi(x) = f_0(x) + \alpha \sum_{i=1}^m \max\{0, f_i(x)\}^2$, where $\alpha > 0$.

a) Show that $\phi(x)$ is convex.

b) Suppose $\tilde{x}$ minimizes $\phi(x)$. Show how to find, from $\tilde{x}$, a feasible point for the dual of the original problem (1).

c) Find the corresponding lower bound on the optimal value of the original problem (1) based on your answer to part (b).

6) Suppose we are given a convex problem

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \cdots, m, \end{array} \quad (2)$$

with the corresponding dual problem $\underset{\lambda\succeq 0}{\text{maximize}}\ g(\lambda)$. In addition, assume that Slater's condition holds, i.e., we have strong duality and the dual optimum is attained. For simplicity also assume that there is a unique dual optimal solution λ^*. A penalty term of the form $f_i(x)^+ = \max\{0, f_i(x)\}$ is defined for (2). Therefore, for fixed $t > 0$, we formulate the following penalized unconstrained problem

$$\text{minimize } f_0(x) + t \max_{i=1,\ldots,m} f_i(x)^+. \quad (3)$$

Show that the defined penalty function is *exact*, i.e., for t large enough, the solution of the unconstrained problem (3) is also a solution of (2):

a) Show that the objective function in (3) is convex.

b) We can express (3) as

$$\begin{aligned} \text{minimize } & f_0(x) + ty \\ \text{subject to } & f_i(x) \le y, \quad i = 1\ldots, m \\ & 0 \le y, \end{aligned}$$

with variables $x, y \in \mathbf{R}$. Find the corresponding Lagrange dual problem and express it in terms of the Lagrange dual function $g(\cdot)$ for problem (2).

c) Use the result in (b) to prove the following property. If $t > \mathbf{1}^T\lambda^*$, then any minimizer of (3) is also an optimal solution of (2).