Homework 3 (Computational)

Due on Thursday, October 11, at 5:20 pm

Solve the following 5 problems to get full credit for this homework assignment. Please follow the instructions for homework assignments. I reserve the right to deduct points if you do not follow these rules. Questions 1 and 2 can be answered with plain Matlab code. Questions 3, 4 and 5 require `cvx` (a user guide can be found here: `http://cvxr.com/cvx/doc/index.html`). The computational part of the homework can be submitted in groups of up to 3 people. Please send me an email (homework.am.math@gmail.com) if you intend to submit the homework with one or two of your fellow students. This information needs to be included in your final submission, also.

1. Let $A \in \mathbf{R}^{n,n}$, $b \in \mathbf{R}^n$. Derive the Hessian for the least squares problem

$$\underset{x}{\text{minimize}}\ \frac{1}{2}\|Ax - b\|_2^2$$

and the regularized least squares problem

$$\underset{x}{\text{minimize}}\ \frac{1}{2}\|Ax - b\|_2^2 + \frac{\beta}{2}\|Bx\|_2^2,$$

with regularization parameter $\beta > 0$. In our initial formulation we assumed

$$B = \text{diag}(1, \ldots, 1) \in \mathbf{R}^{n,n}.$$

Make the current implementation more generic, to allow for arbitrary B (we will use $B = \text{diag}(1, \ldots, 1)$ for now). Extend the implementation of the derivative check based on the Taylor expansion (homework 2) to assess if the Hessian matrix is correct (if not, fix it). How would you expect the error to behave in general, and how does it behave for these problems? Explain your observations (i.e., add comments to your code that explain what you observe). Submit your scripts and functions as described in the general instructions for the homework.

2. The regularized least squares problem

$$\underset{x}{\text{minimize}}\ \frac{1}{2}\|Ax - b\|_2^2 + \frac{\beta}{2}\|Bx\|_2^2$$

allows for an analytical solution. Since the Hessian is positive definite, any stationary point is optimal. Since this is a quadratic, convex problem, a stationary point $x^\star$ will satisfy the optimality condition $g(x^\star) = 0$, where $g(x)$ is the gradient of the objective function. Use this equality to derive a system of the form $Lx = y$. To construct a (synthetic) test problem, use the matrix $A = \tilde{A}^\mathsf{T}\tilde{A} + 1\text{e}{-3}\,\text{diag}(1,1,1) \in \mathbf{R}^{3,3}$, where

$$\tilde{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \quad \text{and} \quad x^\star = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

represents the true solution to our problem. We construct b to be (approximately) in the column space of A. That is, $b = Ax^\star + \delta x$, where $\delta x \in \mathbf{R}^3$ represents noise (i.e., a random perturbation). Use `deltax = 0.01*randn(3,1)` to compute δx. Solve the linear system $Ax = b$ and $Lx = y$ by using Matlab's backslash command, i.e., e.g., `x = A \ b`. Compare the solutions x_{sol} of $Lx = y$ (for $\beta = 1$

and $B = \mathrm{diag}(1,1,1)$) and the solution for $Ax = b$ to the true solution $x^\star$. That is, print the values of x_{sol} and the relative error $\|x^\star - x_{\text{sol}}\|_2 / \|x^\star\|_2$. Add comments to your script that describe what you observe. Submit your scripts and functions as described in the general instructions for the homework.

3. In `ex05_cheb.m` in the `optik` repository (`https://github.com/andreasmang/optik`) I provide an exemplary code that solves the problem of finding the largest Euclidean ball that lies in a polyhedron P described by linear inequalities, i.e.,

$$P := \{x \in \mathbf{R}^n : a_i^\mathsf{T} x \le b_i, \;\; i = 1, \ldots, m\}.$$

We have seen in class that this problem can be formulated as the linear program

$$\begin{array}{ll} \underset{r \in \mathbf{R},\, x_c \in \mathbf{R}^n}{\text{maximize}} & r \\ \text{subject to} & a_i^\mathsf{T} x_c + r\|a_i\|_2 \le b_i, \;\; i = 1, \ldots, m. \end{array}$$

Run the code and try to understand it. Write a function that takes a matrix A and a vector b as input, and generalizes the code to an arbitrary number of inequalities $a_i^\mathsf{T} x_c + r\|a_i\|_2 \le b_i$. The outputs of the function should be x_c and r. Write a second function that generalizes the plot commands. This function takes A, x_c, r, and b as input and visualizes the computed solution. In a first step, use the vectors that are provided in `ex05_cheb.m` to construct A and b to make sure that your generalized implementation works. Then (within the same script) generate another example with two more inequalities, for which there is still a point inside the polyhedron. Submit your scripts and functions as described in the general instructions for the homework.

4. We consider a linear dynamical system with state $x(t) \in \mathbf{R}^n$, $t = 0, ..., n_t$, and actuator or input signal $u(t) \in \mathbf{R}$ for $t = 0, ..., n_t - 1$. The dynamics of the system are given by the linear recurrence

$$x(t+1) = Ax(t) + bu(t), \;\; t = 0, ..., n_t - 1,$$

where $A \in \mathbf{R}^{n,n}$ and $b \in \mathbf{R}^n$ are given. We assume that the initial state is zero, i.e., $x(0) = 0$. The minimum fuel optimal control problem is to choose the inputs $u(0), ..., u(n_t - 1)$ so as to minimize the total fuel consumed, which is given by $f = \sum_{t=0}^{n_t-1} f(u(t))$ subject to the constraint that $x(n_t) = x_{\text{des}}$, where n_t is the (given) time horizon, and $x_{\text{des}} \in \mathbf{R}^n$ is the (given) desired final or target state. The function $f : \mathbf{R} \to \mathbf{R}$ is the fuel use map for the actuator, and gives the amount of fuel used as a function of the actuator signal amplitude. In this problem we use

$$f(a) = \begin{cases} |a| & |a| \le 1 \\ 2|a| - 1 & |a| > 1. \end{cases}$$

This means that fuel use is proportional to the absolute value of the actuator signal, for actuator signals between -1 and 1; for larger actuator signals the marginal fuel efficiency is half. Solve the minimum fuel optimal control problem described above, for the instance with problem data

$$A = \begin{bmatrix} -1 & 0.4 & 0.8 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 0 \\ 0.3 \end{bmatrix}, \quad x_{\text{des}} = \begin{bmatrix} 7 \\ 2 \\ -6 \end{bmatrix},$$

and $n_t = 30$. Establish the basic problem formulation (minimize the total fuel consumed subject to the dynamical system) and solve it using `cvx`. Plot the actuator signal $u(t)$ as a function of time t.

5. We consider the selection of n non-negative activity levels, denoted $x_1, \dots, x_n$. These activities consume m resources, which are limited. Activity j consumes $a_{ij}x_j$ of resource i, where a_{ij} are given. The total resource consumption is additive, so the total of resource i consumed is $c_i = \sum_{j=1}^n a_{ij}x_j$. Ordinarily, we have $a_{ij} \geq 0$, i.e., activity j consumes resource i. But we allow the possibility that $a_{ij} < 0$, which means that activity j actually generates resource i as a by-product. Each resource consumption is limited: We must have $c_i \leq c_{\mathsf{max}}$, where c_{max} are given. Each activity generates revenue, which is a piecewise-linear concave function of the activity level:

$$r_j(x_j) = \begin{cases} p_j x_j & 0 \leq x_j \leq q_j \\ p_j q_j + \tilde{p}_j(x_j - q_j) & x_j \geq q_j. \end{cases}$$

Here, $p_j > 0$ is the basic price, $q_j > 0$ is the quantity discount level, and $\tilde{p}_j$ is the quantity discount price, for (the product of) activity j. (We have $0 < \tilde{p}_j < p_j$.) The total revenue is the sum of the revenues associated with each activity, i.e., $\sum_{j=1}^n r_j(x_j)$. The goal is to choose activity levels that maximize the total revenue while respecting the resource limits. In general, we can reformulate this problem as a linear program. This is not required by `cvx`. Establish the basic problem formulation (maximization of the total revenue subject to inequality constraints) and solve it using `cvx` for the instance with problem data

$$A = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 3 & 1 \\ 0 & 3 & 1 & 1 \\ 2 & 1 & 2 & 5 \\ 1 & 0 & 3 & 2 \end{bmatrix}, \quad c_{\mathsf{max}} = \begin{bmatrix} 100 \\ 100 \\ 100 \\ 100 \\ 100 \end{bmatrix}, \quad p = \begin{bmatrix} 3 \\ 2 \\ 7 \\ 6 \end{bmatrix}, \quad \tilde{p} = \begin{bmatrix} 2 \\ 1 \\ 4 \\ 2 \end{bmatrix}, \quad \text{and} \quad q = \begin{bmatrix} 4 \\ 10 \\ 5 \\ 10 \end{bmatrix}.$$

Give the optimal activity levels, the revenue generated by each one, and the total revenue generated by the optimal solution. Also, give the average price per unit for each activity level, i.e., the ratio of the revenue associated with an activity, to the activity level. (These numbers should be between the basic and discounted prices for each activity.) Give a brief description explaining, or at least commenting on, the solution you find (add this as a comment to your code). Submit your script as described in the general instructions for the homework.