EE 127 3/18/14
L. El Ghaoui

# Homework Assignment #4

**Due date:** 4/15/14, in class.

**Exercise 1 (Boolean least-squares)** Consider the following problem, known as *Boolean Least Squares*:

$$\phi = \min_x \|Ax - b\|_2^2 \ : \ x_i \in \{-1, 1\}, \ \ i = 1, \ldots, n.$$

Here, the variable is $x \in \mathbb{R}^n$, where $A \in \mathbb{R}^{m,n}$ and $b \in \mathbb{R}^m$ are given. This is a basic problem arising, for instance, in digital communications. A brute force solution is to check all $2^n$ possible values of $x$, which is usually impractical.

1. Show that the problem is equivalent to

$$\begin{aligned} \phi = \min_{X,x} \quad & \mathbf{Tr}(A^\top A X) - 2b^\top A x + b^\top b \\ \text{s.t.} \quad & X = xx^\top, \\ & X_{ii} = 1, \quad i = 1, \ldots, n, \end{aligned}$$

   in the variables $X = X^\top \in \mathbb{R}^{n,n}$ and $x \in \mathbb{R}^n$.

2. The constraint $X = xx^\top$, i.e., the set of rank-1 matrices is not convex, therefore the problem is still hard. However, an efficient approximation can be obtained by relaxing this constraint to $X \succeq xx^\top$, obtaining

$$\begin{aligned} \phi \geq \phi_{\text{sdp}} = \min_X \quad & \mathbf{Tr}(A^\top A X) - 2b^\top A x + b^\top b \\ \text{s.t.} \quad & \begin{bmatrix} X & x \\ x^\top & 1 \end{bmatrix} \succeq 0, \\ & X_{ii} = 1, \quad i = 1, \ldots, n. \end{aligned}$$

   The relaxation produces a lower-bound to the original problem. Once that is done, an approximate solution to the original problem can be obtained by rounding the solution: $x_{\text{sdp}} = \text{sgn}(x^*)$, where $x^*$ is the optimal solution of the semidefinite relaxation.

3. Another approximation method is to relax the non-convex constraints $x_i \in \{-1, 1\}$ to convex interval constraints $-1 \leq x_i \leq 1$ for all $i$, which can be written $\|x\|_\infty \leq 1$. Therefore a different lower bound is given by:

$$\phi \geq \phi_{\text{int}} \doteq \min \|Ax - b\|_2^2 \ : \ \|x\|_\infty \leq 1.$$

   Once that problem is solved, we can round the solution by $x_{\text{int}} = \text{sgn}\,(x^*)$ and compare the original objective value $\|Ax_{\text{int}} - b\|_2^2$.

4. Which one of $\phi_{\rm sdp}$ and $\phi_{\rm int}$ produces the closest approximation to $\phi$? Justify carefully your answer.

5. Use now 100 independent realizations with normally distributed data, $A \in \mathbb{R}^{10,10}$ (independent entries with mean zero) and $b \in \mathbb{R}^{10}$ (independent entries with mean 1). Plot and compare the histograms of $\|Ax_{\rm sdp} - b\|_2^2$ of part 2, $\|Ax_{\rm int} - b\|_2^2$ of part 3, and the objective corresponding to a naïve method $\|Ax_{\rm ls} - b\|_2^2$, where $x_{\rm ls} = \mathrm{sgn}\left((A^\top A)^{-1}A^\top b\right)$ is the rounded ordinary Least Squares solution. Briefly discuss accuracy and computation time (in seconds) of the three methods.

6. Assume that, for some problem instance, the optimal solution $(x, X)$ found via the SDP approximation is such that $x$ belongs to the original non-convex constraint set $\{x : x_i \in \{-1, 1\},\ \ i = 1, \ldots, n\}$. What can you say about the SDP approximation in that case?

**Exercise 2 (Non-negativity of polynomials)** A second-degree polynomial with values $p(x) = y_0 + y_1 x + y_2 x^2$ is non-negative everywhere if and only if

$$\forall\, x\ :\ \begin{bmatrix} 1 \\ x \end{bmatrix}^\top \begin{bmatrix} y_0 & y_1/2 \\ y_1/2 & y_2 \end{bmatrix} \begin{bmatrix} 1 \\ x \end{bmatrix} \geq 0,$$

which in turn can be written as an LMI in $y = (y_0, y_1, y_2)$:

$$\begin{bmatrix} y_0 & y_1/2 \\ y_1/2 & y_2 \end{bmatrix} \succeq 0.$$

In this exercise, you show a more general result, which applies to any polynomial of even degree $2k$ (polynomials of odd degree can't be non-negative everywhere). To simplify, we only examine the case $k = 2$, that is, fourth-degree polynomials; the method employed here can be generalized to $k > 2$.

1. Show that a fourth-degree polynomial $p$ is non-negative everywhere if and only if it is a sum of squares, that is, it can be written as

$$p(x) = \sum_{i=1}^{4} q_i(x)^2,$$

where $q_i$'s are polynomials of degree at most two. *Hint:* show that $p$ is non-negative everywhere if and only if it is of the form

$$p(x) = p_0 \left((x - a_1)^2 + b_1^2\right) \left((x - a_2)^2 + b_2^2\right),$$

for some appropriate real numbers $a_i, b_i$, $i = 1, 2$, and some $p_0 \geq 0$.

2. Using the previous part, show that if a fourth-degree polynomial is a sum of squares, then it can be written as

$$p(x) = \begin{bmatrix} 1 & x & x^2 \end{bmatrix} Q \begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix}. \tag{1}$$

for some positive-semidefinite matrix $Q$.

3. Show the converse: if a positive semi-definite matrix $Q$ satisfies condition (1) for every $x$, then $p$ is a sum of squares. *Hint:* use a factorization of $Q$ of the form $Q = AA^\top$, for some appropriate matrix $A$.

4. Show that a fourth-degree polynomial $p(x) = y_0 + y_1 x + y_2 x^2 + y_3 x^3 + y_4 x^4$ is non-negative everywhere if and only if there exist a $3 \times 3$ matrix $Q$ such that

$$Q \succeq 0, \quad y_l = \sum_{i+j=l} Q_{i+1,j+1}, \quad l = 0, \ldots, 4.$$

*Hint:* equate the coefficients of the powers of $x$ in the left and right sides of equation (1).

**Exercise 3 (Design of a water reservoir)** We need to design a water reservoir for water and energy storage, as depicted in Figure 1.

The concrete basement has square section of side length $b_1$ and height $h_0$, while the reservoir itself has square section of side length $b_2$ and height $h$. Some useful data is reported in Table 1.

The critical load limit $N_{cr}$ of the basement should withstand at least twice the weight of water. The structural specification $h_0/b_1^2 \leq 35$ should hold. The form factor of the reservoir should be such that $1 \leq b_2/h \leq 2$. The total height of the structure should be no larger than 30 m. The total weight of the structure (basement plus reservoir full of water) should not exceed $9.8 \times 10^5$ N. The problem is to find the dimensions $b_1, b_2, h_0, h$ such that the potential energy $P_w$ of the stored water is maximal (assume $P_w = (\rho_w h b_2^2) h_0$). Explain if and how the problem can be modeled as a convex optimization problem and, in the positive case, find the optimal design.

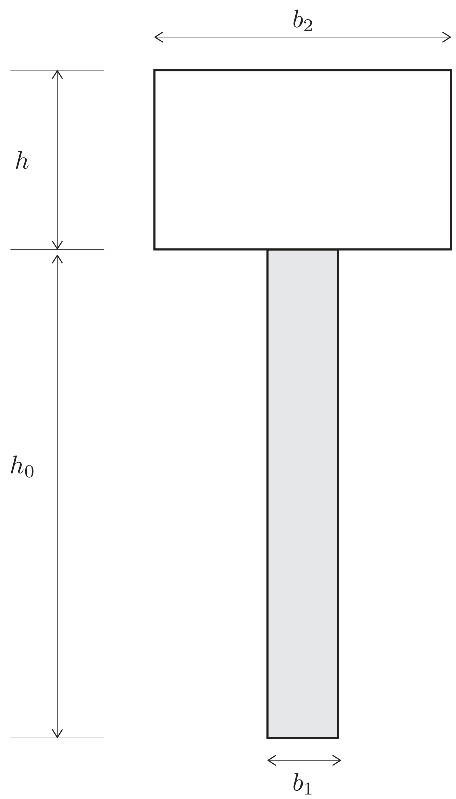


Figure 1: A water reservoir on concrete basement.

| Quantity | Value | Units | Description |
|---|---|---|---|
| $g$ | 9.8 | $m/s^2$ | gravity acceleration |
| $E$ | $30 \times 10^9$ | $N/m^2$ | basement long. elasticity modulus |
| $\rho_w$ | $10 \times 10^3$ | $N/m^3$ | specific weight of water |
| $\rho_b$ | $25 \times 10^3$ | $N/m^3$ | specific weight of basement |
| $J$ | $b_1^4/12$ | $m^4$ | basement moment of inertia |
| $N_{cr}$ | $\pi^2 JE/(2h_0)^2$ | N | basement critical load limit |

Table 1: Data for reservoir problem.

**Exercise 4 (Image deformation)** A rigid transformation is a mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$ that is the composition of a translation and a rotation. Mathematically, we can express a rigid transformation $\phi$ as $\phi(x) = Rx + r$, where $R$ is an $n \times n$ orthogonal transformation and $r \in \mathbb{R}^n$ a vector.

We are given a set of pairs of points $(x_i, y_i)$ in $\mathbb{R}^n$, $i = 1, \ldots, m$, and wish to find a rigid transformation that best matches them. We can write the problem as

$$\min_{R \in \mathbb{R}^{n,n},\, r \in \mathbb{R}^n} \sum_{i=1}^{m} \|Rx_i + r - y_i\|_2^2 \;:\; R^\top R = I_n, \tag{2}$$

where $I_n$ is the $n \times n$ identity matrix.

The problem arises in image processing, to provide ways to deform an image (represented as a set of two-dimensional points) based on the manual selection of a few points and their transformed counterparts.

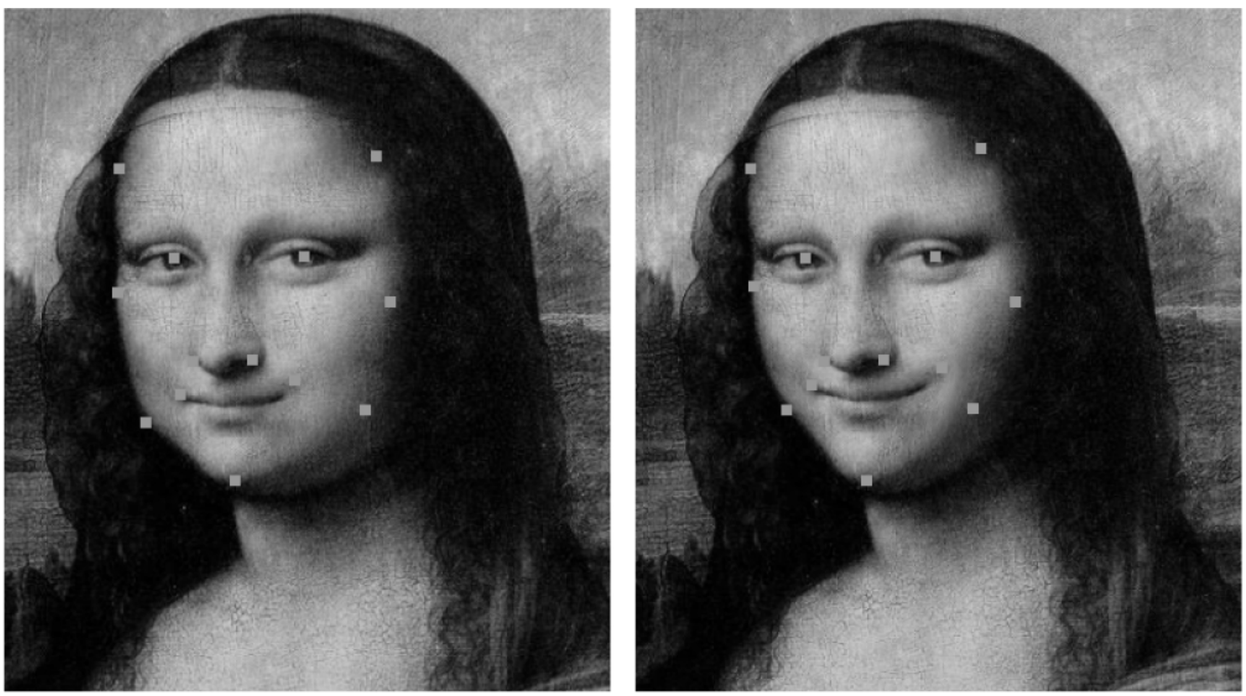

Figure 2: Image deformation via rigid transformation. The image on the left is the original image, and that on the right is the deformed image. Yellow dots indicate points for which the deformation is chosen by the user.

1. Assume that $R$ is fixed in problem (2). Express an optimal $r$ as a function of $R$.

2. Show that the corresponding optimal value (now a function of $R$ only) writes as the original objective function, with $r = 0$ and $x_i, y_i$ replaced with their centered counterparts,

$$\bar{x}_i = x_i - \hat{x}, \quad \hat{x} = \frac{1}{m}\sum_{j=1}^{m} x_j, \quad \bar{y}_i = y_i - \hat{y}, \quad \hat{y} = \frac{1}{m}\sum_{j=1}^{m} y_j.$$

3. Show that the problem can be written as

$$\min_R \ \|RX - Y\|_F \ : \ R^\top R = I_n,$$

for appropriate matrices $X, Y$, which you will determine. *Hint:* explain why you can square the objective; then expand.

4. Show that the problem can be further written as

$$\max_R \ \mathbf{Tr} RZ \ : \ R^\top R = I_n,$$

for an appropriate $n \times n$ matrix $Z$, which you will determine.

5. Show that $R = VU^\top$ is optimal, where $Z = USV^\top$ is the SVD of $Z$. *Hint:* reduce the problem to the case when $Z$ is diagonal, and use without proof the fact that when $Z$ is diagonal, $I_n$ is optimal for the problem.

6. Show the result you used in the previous question: assume $Z$ is diagonal, and show that $R = I_n$ is optimal for the problem above. *Hint:* show that $R^\top R = I_n$ implies $|R_{ii}| \leq 1$, $i = 1, \ldots, n$, and using that fact, prove that the optimal value is less than or equal to $\mathbf{Tr} Z$.

7. How woud you apply this technique to make Mona Lisa smile more? *Hint:* in Figure 2, the two-dimensional points $x_i$ are given (as yellow dots) on the left panel, while the corresponding points $y_i$ are shown on the left panel. These points are manually selected. The problem is to find how to transform all the other points in the original image.