Homework assignment 1

1. Suppose A is an $m \times n$ matrix with first column $\mathbf{1}$ (the m-vector of ones) and b is an m-vector. Let $\hat{x}$ be a solution of the least squares problem

$$\text{minimize} \quad \|Ax - b\|_2^2$$

and define $\hat{r} = b - A\hat{x}$. Use the normal equations $A^T(A\hat{x} - b) = 0$ to show that

$$\mathbf{avg}(A\hat{x}) = \mathbf{avg}(b), \qquad \mathbf{std}(A\hat{x})^2 + \mathbf{std}(\hat{r})^2 = \mathbf{std}(b)^2$$

where, for an m-vector y,

$$\mathbf{avg}(y) = \frac{1}{m}(\mathbf{1}^T y), \qquad \mathbf{std}(y) = \frac{1}{\sqrt{m}}\|y - \mathbf{avg}(y)\mathbf{1}\|_2.$$

(These quantities are known as the average and the standard deviation of the vector.)

2. Let X be a symmetric matrix partitioned as

$$X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}. \tag{1}$$

If A is nonsingular, the matrix $S = C - B^T A^{-1} B$ is called the *Schur complement* of A in X. It can be shown that if A is positive definite, then $X \succeq 0$ (X is positive semidefinite) if and only if $S \succeq 0$ (see page 650 of the textbook). In this exercise we prove the extension of this result to singular A mentioned on page 651 of the textbook.

(a) Suppose $A = 0$ in (1). Show that $X \succeq 0$ if and only if $B = 0$ and $C \succeq 0$.

(b) Let A be a symmetric $n \times n$ matrix with eigenvalue decomposition

$$A = Q\Lambda Q^T,$$

where Q is orthogonal ($Q^TQ = QQ^T = I$) and $\Lambda = \mathbf{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Assume the first r eigenvalues λ_i are nonzero and $\lambda_{r+1} = \cdots = \lambda_n = 0$. Partition Q and Λ as

$$Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}, \qquad \Lambda = \begin{bmatrix} \Lambda_1 & 0 \\ 0 & 0 \end{bmatrix}$$

with Q_1 of size $n \times r$, Q_2 of size $n \times (n-r)$, and $\Lambda_1 = \mathbf{diag}(\lambda_1, \ldots, \lambda_r)$. The matrix

$$A^\dagger = Q_1 \Lambda_1^{-1} Q_1^T$$

is called the *pseudo-inverse* of A. Verify that

$$AA^\dagger = A^\dagger A = Q_1 Q_1^T, \qquad I - AA^\dagger = I - A^\dagger A = Q_2 Q_2^T.$$

The matrix-vector product $AA^\dagger x = Q_1 Q_1^T x$ is the orthogonal projection of the vector x on the range of A. The matrix-vector product $(I - AA^\dagger)x = Q_2 Q_2^T x$ is the projection on the nullspace.

(c) Show that the block matrix X in (1) is positive semidefinite if and only if

$$A \succeq 0, \qquad (I - AA^\dagger)B = 0, \qquad C - B^T A^\dagger B \succeq 0.$$

(The second condition means that the columns of B are in the range of A.)

Hint. Let $A = Q\Lambda Q^T$ be the eigenvalue decomposition of A. The matrix X in (1) is positive semidefinite if and only if the matrix

$$\begin{bmatrix} Q^T & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} \Lambda & Q^T B \\ B^T Q & C \end{bmatrix}$$

is positive semidefinite. Use the observation in part (a) and the Schur complement characterization for positive definite 2×2 block matrices to show the result.

3. This problem is an introduction to the MATLAB software package CVX that will be used in the course. CVX can be downloaded from `www.cvxr.com`.

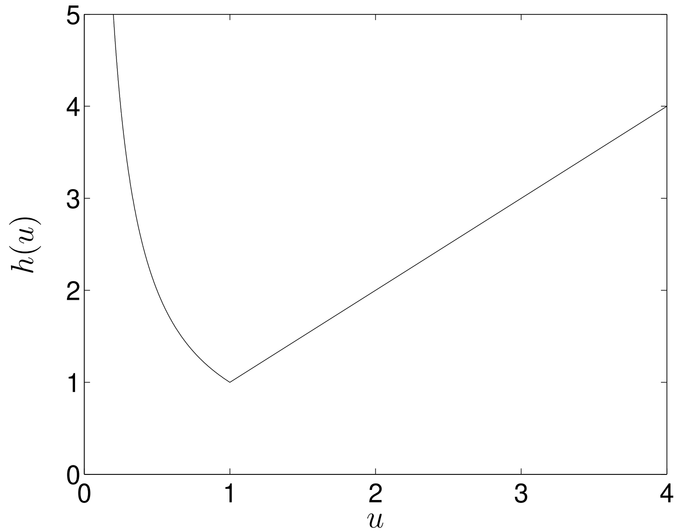
We consider the illumination problem of lecture 1. We take $I_{\text{des}} = 1$ and $p_{\max} = 1$, so the problem is

$$\begin{array}{ll} \text{minimize} & f_0(p) = \max\limits_{k=1,\ldots,n} |\log(a_k^T p)| \\ \text{subject to} & 0 \le p_j \le 1, \quad j = 1, \ldots, m, \end{array} \tag{2}$$

with variable $p \in \mathbf{R}^m$. As mentioned in the lecture, the problem is equivalent to

$$\begin{array}{ll} \text{minimize} & \max\limits_{k=1,\ldots,n} h(a_k^T p) \\ \text{subject to} & 0 \le p_j \le 1, \quad j = 1, \ldots, m, \end{array} \tag{3}$$

where $h(u) = \max\{u, 1/u\}$ for $u > 0$. The function h, shown in the figure below, is nonlinear, nondifferentiable, and convex.



To see the equivalence between (2) and (3), we note that

$$\begin{aligned} f_0(p) &= \max_{k=1,\ldots,n} |\log(a_k^T p)| \\ &= \max_{k=1,\ldots,n} \max\{\log(a_k^T p), \log(1/a_k^T p)\} \\ &= \log \max_{k=1,\ldots,n} \max\{a_k^T p, 1/a_k^T p\} \\ &= \log \max_{k=1,\ldots,n} h(a_k^T p), \end{aligned}$$

and since the logarithm is a monotonically increasing function, minimizing f_0 is equivalent to minimizing $\max_{k=1,\ldots,n} h(a_k^T p)$.

The problem data are given in the file `illum_data.m` posted on the course website. Executing this file in MATLAB creates the $n \times m$-matrix A (which has rows a_k^T). There are 10 lamps ($m = 10$) and 20 patches ($n = 20$).

Use the following methods to compute four approximate solutions and the exact solution, and compare the answers (the vectors p and the corresponding values of $f_0(p)$).

(a) *Equal lamp powers.* Take $p_j = \gamma$ for $j = 1, \ldots, m$. Plot $f_0(p)$ versus γ over the interval $[0, 1]$. Graphically determine the optimal value of γ, and the associated objective value. The objective function $f_0(p)$ can be evaluated in MATLAB as `max(abs(log(A*p)))`.

(b) *Least-squares with saturation.* Solve the least squares problem

$$\text{minimize} \quad \sum_{k=1}^{n} (a_k^T p - 1)^2 = \|Ap - \mathbf{1}\|_2^2.$$

If the solution has negative coefficients, set them to zero; if some coefficients are greater than 1, set them to 1. Use the MATLAB command `x = A \ b` to solve a least squares problem (minimize $\|Ax - b\|_2^2$).

(c) *Regularized least squares.* Solve the regularized least squares problem

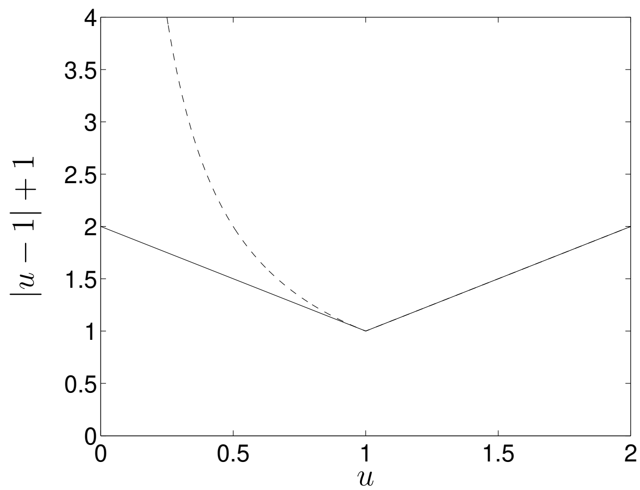
$$\text{minimize} \quad \sum_{k=1}^{n}(a_k^T p - 1)^2 + \rho \sum_{j=1}^{m}(p_j - 0.5)^2 = \|Ap - \mathbf{1}\|_2^2 + \rho \|p - (1/2)\mathbf{1}\|_2^2,$$

where $\rho > 0$ is a parameter. Increase ρ until all coefficients of p are in the interval $[0, 1]$.

(d) *Chebyshev approximation.* Solve the problem

$$\begin{array}{ll} \text{minimize} & \max_{k=1,\ldots,n} |a_k^T p - 1| = \|Ap - \mathbf{1}\|_\infty \\ \text{subject to} & 0 \le p_j \le 1, \quad j = 1, \ldots, m. \end{array}$$

We can think of this problem as obtained by approximating the nonlinear function $h(u)$ by a piecewise-linear function $|u-1|+1$. As shown in the figure below, this is a good approximation around $u = 1$. This problem can be converted to a linear program and solved using the MATLAB function `linprog`. It can also be solved directly in CVX, using the expression `norm(A*p - 1, inf)` to specify the cost function.



(e) *Exact solution.* Finally, use CVX to solve

$$\begin{array}{ll} \text{minimize} & \max_{k=1,\ldots,n} \max(a_k^T p, 1/a_k^T p) \\ \text{subject to} & 0 \le p_j \le 1, \quad j = 1, \ldots, m. \end{array}$$

Use the CVX function `inv_pos()` to expresss the function $f(x) = 1/x$ with domain $\mathbf{R}_{++}$.

4. Exercise T2.7. (Exercise 2.7 in the textbook.)

5. Exercise T2.12(g).