Optimization 1 — Homework 3

November 5, 2020

Problem 1
The principal minors criterion states that a symmetric matrix $\mathbf{A}$ is PD if and only if all the leading principal minors of $\mathbf{A}$ are PD. The purpose of this problem is to give a proof of this result using optimization techniques.
Let $\mathbf{A} \in \mathbb{R}^{(n+1)\times(n+1)}$ and $\mathbf{B} \in \mathbb{R}^{n\times n}$ be two symmetric matrices such that

$$\mathbf{A} = \begin{bmatrix} \mathbf{B} & \mathbf{b} \\ \mathbf{b}^T & c \end{bmatrix},$$

where $\mathbf{B} \succ 0$, $\mathbf{b} \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

(a) Consider the function $p\colon \mathbb{R}^n \to \mathbb{R}$ defined by

$$p(\mathbf{x}) = \begin{pmatrix} \mathbf{x}^T & 1 \end{pmatrix} \mathbf{A} \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix}.$$

Show that $\mathbf{x}^* = -\mathbf{B}^{-1}\mathbf{b}$ is the unique global minimizer of p over $\mathbb{R}^n$ and that p is positive over its domain if and only if $c > \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b}$.

(b) Show that $\det(\mathbf{A}) = \det(\mathbf{B}) \cdot \left(c - \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b}\right)$.

Hint: Find a vector $\mathbf{d} \in \mathbb{R}^n$ such that

$$\begin{bmatrix} \mathbf{I} & \mathbf{d} \\ \mathbf{0} & 1 \end{bmatrix}^T \mathbf{A} \begin{bmatrix} \mathbf{I} & \mathbf{d} \\ \mathbf{0} & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & c - \mathbf{b}^T\mathbf{B}^{-1}\mathbf{b} \end{bmatrix}.$$

(c) Prove the principal minors criterion by induction on the dimension of $\mathbf{A}$.

Problem 2
Assume that each point in the set $\mathcal{X} = \{(x_i, y_i)\}_{i=1}^m \subset \mathbb{R}^2$ can be approximated by the relation

$$y = \frac{c_0 + c_1 x + c_2 x^2}{d_0 + d_1 x + d_2 x^2},$$

where $\mathbf{u} = (c_0, c_1, c_2, d_0, d_1, d_2)^T \in \mathbb{R}^6$ is an unknown vector of coefficients. Since for all scalar $\alpha \neq 0$ the coefficients $\mathbf{u}$ and $\alpha\mathbf{u}$ represent the same curve, the above relation has redundancy (meaning, there are infinitely many curves).

(a) One approach to tackle the redundancy is to assume that $d_0 = 1$. Write a least squares problem that finds an approximating vector $\mathbf{u}$.

(b) Write a MATLAB function that is called by

```
[c0,c1,c2,d1,d2]=fit_rational(X)
```

where `X` is a matrix with m rows and two columns that contains the set $\mathcal{X}$. The function solves your least squares problem from section (a). Use the MATLAB operator \. Run the function on the file `.curve` that is in the moodle site. Write the output coefficients and plot the resulting curve with the given points.

(c) Another approach to tackle the redundancy is to assume that $\left\| (c_0, c_1, c_2, d_0, d_1, d_2)^T \right\| = 1$. Write an optimization problem with a quadratic objective function that finds an approximating vector $\mathbf{u}$. Explain how you would solve this problem (you can use the Rayleigh quotient as seen in class in your explanation).

(d) Write a MATLAB function that is called by

```
[c0,c1,c2,d0,d1,d2]=fit_rational_normed(X)
```

that solves your optimization problem from section (c). Which approach yields a better approximation?

Problem 3

For each of the following functions, determine whether it is coercive or not:

(a) $f(x, y) = (x - 2y)^4 + 64xy$.

(b) $f(x, y) = \begin{cases} x^2 - 2xy + y^2, & x \neq y, \\ x^2 + y^2, & x = y. \end{cases}$

(c) $f(\mathbf{x}) = \frac{\mathbf{x}^T \mathbf{A} \mathbf{x}}{\|\mathbf{x}\|+1}$ where $\mathbf{A} \in \mathbb{R}^{n \times n}$ is PD.

Problem 4

Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be defined as $f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} + 2\mathbf{b}^T \mathbf{x} + c$, where $\mathbf{A} \in \mathbb{R}^{n \times n}$ is symmetric, $\mathbf{b} \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Suppose that $\mathbf{A} \succeq 0$. Show that f is bounded from below over $\mathbb{R}^n$ if and only if $\mathbf{b} \in \text{Image}(\mathbf{A})$.

Problem 5

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{L} \in \mathbb{R}^{p \times n}$and $\lambda \in \mathbb{R}_{++}$. Consider the function

$$f(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 + \lambda \|\mathbf{L}\mathbf{x}\|^2 .$$

Show that f has a unique minimum if and only if $\text{Ker}(\mathbf{A}) \cap \text{Ker}(\mathbf{L}) = \{\mathbf{0}_n\}$.