Princeton University

ORF 523
Final Exam, Spring 2020

Tuesday, May 5, 10am EST, to Friday, May 15, 10am EST

Instructor:
A.A. Ahmadi

AIs:
Bachir El Khadir, Jeffrey Zhang,
Emmanuel Ekwedike (grader)

1. Please write out and sign the following pledge on top of the first page of your exam:
   "I pledge my honor that I have not violated the Honor Code or the rules specified by the instructor during this exam."

2. The exam is not to be discussed with *anyone* except possibly the professor and the AIs. You can only ask clarification questions as *public* (and preferably non-anonymous) questions on Piazza. No emails.

3. You are allowed to consult the lecture notes and videos, your own notes, the problem sets and their solutions (yours and ours), the midterm exams and their solutions (yours and ours), the practice exams and their solutions, past Piazza questions and answers, but *nothing else*. You can only use the Internet in case you run into problems related to software. (There should be no need for that either hopefully.)

4. For all problems involving a coding element, show your code. The output that you present should come from your code.

5. The exam is to be submitted on Blackboard before Friday, May 15, at 10 AM EST.

6. Some problems might be harder than others and there is no particular order. Please be rigorous, brief, and to the point in your answers. Good luck!

Problem 1: What is the probability that Zoom's stock goes bust?

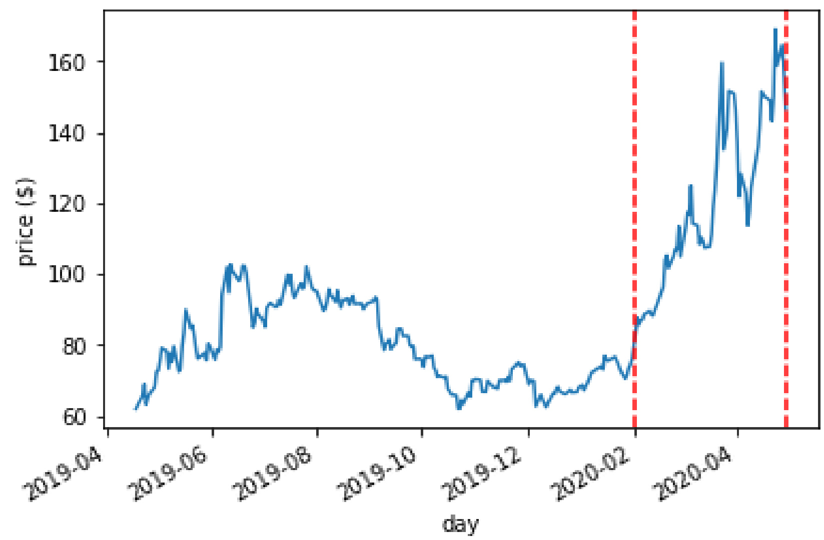

Figure 1: Zoom's stock price

You have noticed that Zoom has been growing in popularity recently (see Figure 1), and you wonder whether you should buy their stock. You download daily stock prices for the duration of three months starting from February 1, 2020 (excluding non-trading days) and compute the daily returns as

$$r_i = \frac{P_i - P_{i-1}}{P_i}, \quad i = 1, \ldots, 61,$$

where P_i is the price of the stock on day i. You assume that the daily returns r_i are independent copies of a random variable r with unknown distribution supported on the interval $[-0.4, 0.4]$ (i.e., the daily returns never fall below -40% or go above 40%). From the data and for $k = 1, \ldots, 4$, you compute the empirical means m_k of the k-th moment $\mathbb{E}[r^k]$ of r:

$$m_1 = 0.0068, m_2 = 0.0034, m_3 = 2 \times 10^{-6}, m_4 = 5 \times 10^{-5}. \tag{1}$$

Given that you are risk averse, you decide that you should buy Zoom's stock only if the probability that daily returns go below -0.1 is small. The problem, however, is that you do not know how to compute this probability as you don't know the distribution of the daily returns. You decide instead to compute the *worst-case* probability over all distributions whose first 4 moments are within 10% of those you have computed from data.

1. Let

$$
\begin{aligned}
\alpha := \inf_{q,r,s,\gamma} \quad & \gamma \\
\text{s.t.} \quad & q(x) = \sum_{k=0}^{4} q_k x^k \text{ is a degree-4 (univariate) polynomial,} \\
& r(x), s(x) \text{ are quadratic polynomials that are sos,} \\
& q_0 + \sum_{k=1}^{4} q_k m'_k \leq \gamma \ \forall m'_k \in [0.9\ m_k, 1.1\ m_k] \text{ for } k = 1, \dots, 4, \\
& q(x) - \left(0.4^2 - x^2\right) s(x) \text{ is sos,} \\
& q(x) - 1 - (0.4 + x)(-0.1 - x) r(x) \text{ is sos.}
\end{aligned}
$$

Show that

$$\mathbb{P}(r \in [-0.4, -0.1]) \leq \alpha,$$

if the probability is calculated with respect to any distributions on r whose first 4 moments are within 10% of your empirical moments in (1).

Hint: Use the basic fact that for any interval $[a, b]$, $\mathbb{P}(r \in [a, b]) = \mathbb{E}[1_{[a,b]}(r)]$, *where* $1_{[a,b]}(r)$ *is equal to* 1 *if* $r \in [a, b]$ *and* 0 *otherwise.*

2. Compute α to 4 digits after the decimal point.

3. You wonder if the bound α you got from the above problem is overly pessimistic. Find a discrete distribution of returns (i.e., points $x_1, \dots, x_N \in [-0.4, 0.4]$ and probabilities $p_1, \dots, p_N \in [0, 1]$ summing to one) such that

   i) The moments of your discrete distribution are within 10% of the empirical moments of r, i.e., $|\sum_{i=1}^{N} p_i x_i^k - m_k| \leq \frac{m_k}{10}, \quad k = 1, \dots, 4,$

   ii) The probability assigned by your discrete distribution to the interval $[-0.4, -0.1]$ is equal to α; i.e.,

$$\sum_{i \in I} p_i = \alpha, \text{ where } I = \{i \in \{1, \dots, N\} \mid x_i \in [-0.4, -0.1]\}.$$ [1]

   Hint: If q^* *is the quartic polynomial that your solver returns for part 1, a plot of* $q^* - 1_{[-0.4,-0.1]}$ *can help you find the points* $x_1, \dots, x_N$.

[1] To avoid numerical issues, any discrete distribution that assigns to the interval $[-0.4, -0.1]$ a probability larger or equal than 0.99α is acceptable.

Problem 2: Complexity aspects of optimality conditions

1. Consider the decision problem CRITICAL-d:
   Given a polynomial $p : \mathbb{R}^n \to \mathbb{R}$ of degree d with rational coefficients, decide if it has a critical point, i.e., a point $\bar{x} \in \mathbb{R}^n$ such that $\nabla p(\bar{x}) = 0$.

   Show that CRITICAL-d is in P if $d < 3$ and NP-hard if $d \geq 3$.
   Hint: Observe that when $d = 3$, the condition $\nabla p(x) = 0$ gives a set of quadratic equations, but not an arbitrary one! See if you can get around this issue by introducing new variables.

2. What is the largest value of d for which the following statement is true for any polynomial $p : \mathbb{R}^n \to \mathbb{R}$ of degree d and any point $\bar{x} \in \mathbb{R}^n$?

   $$\nabla p(\bar{x}) = 0, \nabla^2 p(\bar{x}) \succ 0 \iff \bar{x} \text{ is a strict local minimum for } p.$$

   Fully justify your answer.

3. Consider the decision problem STRICT-LOCAL-d:
   Given a polynomial $p : \mathbb{R}^n \to \mathbb{R}$ of degree d with rational coefficients and a point $\bar{x} \in \mathbb{Q}^n$, decide if $\bar{x}$ is a strict local minimum for p.

   Show that STRICT-LOCAL-d is in P if $d = 1, 2,$ or 3. (We have seen in class that STRICT-LOCAL-4 is NP-hard.)
   Hint: You can use the fact that the determinant of a matrix with rational entries can be computed in time polynomial in the bit size of the entries of the matrix.

Problem 3: Best subset selection in penalized and constrained forms[2]

Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$ be fixed. For parameters $\lambda > 0$ and $\delta > 0$, consider the following two popular optimization problems in modern statistics which aim for a sparse approximate solution to a system of linear equations:

$$P_\lambda : \quad \inf_{x\in\mathbb{R}^n} \quad \|Ax - b\|_2^2 + \lambda\|x\|_0, \qquad\qquad C_\delta : \quad \begin{array}{ll} \inf\limits_{x\in\mathbb{R}^n} & \|Ax - b\|_2^2 \\ \text{s.t.} & \|x\|_0 \leq \delta. \end{array}$$

Here $\|\cdot\|_0$ denotes the l_0 pseudo-norm of a vector, i.e., the number of its nonzero entries. Let $\Omega(P_\lambda)$ (resp. $\Omega(C_\delta)$) denote the set of optimal solutions of problem P_λ (resp. C_δ).

True or False? If "True", provide a proof. If "False", provide a counterexample and justify why your counterexample is valid.

1. $\forall\lambda > 0$, $\Omega(P_\lambda)$ is nonempty.
2. $\forall\delta > 0$, $\Omega(C_\delta)$ is nonempty.
3. $\forall\lambda > 0, \exists\delta > 0$ such that $\Omega(P_\lambda) = \Omega(C_\delta)$.
4. $\forall\lambda > 0, \forall x \in \Omega(P_\lambda), \exists\delta > 0$ such that $x \in \Omega(C_\delta)$.
5. $\forall\delta > 0, \exists\lambda > 0$ such that $\Omega(C_\delta) = \Omega(P_\lambda)$.
6. $\forall\delta > 0, \forall x \in \Omega(C_\delta), \exists\lambda > 0$ such that $x \in \Omega(P_\lambda)$.

[2] We thank Sinem Uysal for suggesting this problem.

Problem 4: Accounting for nonlinearity and modeling error in stability analysis It is common in control theory to approximate an unknown dynamical system with a linear model, but also to account for nonlinear effects by adding a bounded unknown nonlinear term. More precisely, the dynamics is modelled as

$$x_{k+1} = Ax_k + g(x_k), \tag{2}$$

where $A \in \mathbb{R}^{n\times n}$ is a fixed and $g : \mathbb{R}^n \to \mathbb{R}^n$ is an unknown continuous function satisfying

$$\|g(x)\| \leq \gamma\|x\| \ \forall x \in \mathbb{R}^n \tag{3}$$

for some fixed scalar $\gamma > 0$. An important problem in control is to check whether $x = 0$ is a globally asymptotically stable equilibrium point of the dynamics in (2) for any choice of the function g verifying (3). In order to check this property, one can search for a (homogeneous and coercive) quadratic Lyapunov function $V : \mathbb{R}^n \to \mathbb{R}$ satisfying

$$V(Ax + g(x)) < V(x) \quad \forall x \neq 0, \text{ and for any function } g \text{ verifying (3).} \tag{4}$$

1. Formulate the search for such a Lyapunov function as an SDP feasibility problem.

2. A series of chemical reactions

   $$C_1 \to C_2 \to C_3 \searrow$$

   between three chemical compounds C_1, C_2, and C_3 can be modeled by a dynamical system of the type in (2), where x_k is a 3×1 vector whose i^{th} component $x_{k,i}$ represents the concentration of chemical compound i at time k. Here, the matrix A is given by

   $$\frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix},$$

   representing that at each time step, half of C_1 converts to C_2, half of C_2 converts to C_3, and half of C_3 vanishes. The function $g : \mathbb{R}^3 \to \mathbb{R}^3$ is unknown and represents the hard-to-model nonlinear interactions between the chemical compounds.

   What is the largest value of γ (to two digits after the decimal point) such that if

   $$||g(x)|| \leq \gamma\|x\| \ \forall x \in \mathbb{R}^3,$$

   then all chemical concentrations go to zero irrespective of their initial concentrations?

   Hint: To find lower (resp. upper) bounds on this critical value of γ, leverage part 1 (resp. focus on functions g that are linear).

Problem 5: Convex optimization applied to nonconvex problems

Consider the optimization problem

$$\begin{aligned} \inf_{x\in\mathbb{R}^n} \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \le 0,\ i = 1, \ldots, m, \end{aligned} \tag{5}$$

whose feasible set is nonempty and compact. Suppose for $i = 0, \ldots, m$, the functions $f_i : \mathbb{R}^n \to \mathbb{R}$ can be written as $f_i(x) = g_i(x) - h_i(x)$, where $g_i : \mathbb{R}^n \to \mathbb{R}$ and $h_i : \mathbb{R}^n \to \mathbb{R}$ are all convex functions, h_i's are all differentiable, and g_0 is a strictly convex function. Consider the following algorithm for approximately solving (5):

Algorithm 1

Input: the functions $g_i(x), h_i(x)$ for $i = 0, \ldots, m$, a vector $x_0 \in \mathbb{R}^n$ which is feasible to (5), a positive integer N.

Output: a vector $x_N \in \mathbb{R}^n$.

procedure

  $k \leftarrow 0$

  while $k < N$ **do**

    Let $f_i^k(x) := g_i(x) - \big(h_i(x_k) + \nabla h_i(x_k)^T(x - x_k)\big)$, $i = 0, \ldots, m$

    Solve the optimization problem: $\inf_{x\in\mathbb{R}^n} f_0^k(x)$, s.t. $f_i^k(x) \le 0$, $i = 1, \ldots, m$

    Let x_{k+1} denote its optimal solution

    $k \leftarrow k + 1$

  end while

---

1. Show that the optimization problem solved in each iteration of Algorithm 1 is a convex optimization problem and has a unique optimal solution.
   Hint: You can use the fact that (globally) convex functions are continuous.

2. *Preserving feasibility.* Show that the points $x_1, \ldots, x_N$ generated by Algorithm 1 are all feasible to (5).

3. *The descent property.* Show that the points $x_1, \ldots, x_N$ generated by Algorithm 1 satisfy $f_0(x_{k+1}) \le f_0(x_k)$ for $k = 0, \ldots, N - 1$.

4. Show that any nonconstant polynomial f can be written as $f(x) = g(x) - h(x)$, where g and h are strictly convex polynomials whose degree is at most one higher than f. (Hence, when $f_0, \ldots, f_m$ in (5) are polynomials, Algorithm 1 is applicable.)