

## Homework 3

Out: *Oct 2*Due: *Oct 23***Instructions:**

- Upload your solutions (to the non-extra-credit) to each problem as a **separate PDF** file (one PDF per problem) to codePost. Please make sure you are uploading the correct PDF! Please anonymize your submission (i.e., do not list your name in the PDF), but if you forget, it's OK.
- If you choose to do extra credit, upload your solution to the extra credits as a single separate PDF file to codePost. Please again anonymize your submission.
- You may collaborate with any classmates, textbooks, the Internet, etc. Please upload a brief “collaboration statement” listing any collaborators as a separate PDF on codePost (if you forget, it's OK). But always **write up your solutions individually**.
- For each problem, you should aim to keep your writeup below one page. For some problems, this may be infeasible, and for some problems you may write significantly less than a page. This is not a hard constraint, but part of the assignment is figuring out how to easily convince the grader of correctness, and to do so concisely. “One page” is just a guideline: if your solution is longer because you chose to use figures (or large margins, display math, etc.) that's fine.
- Each problem is worth twenty points (even those with multiple subparts), unless explicitly stated otherwise.

**Problems:**

- §1 Review the **Portfolio Management** case study from the Lecture 16 notes and the Constant Rebalanced Portfolio (CRP) strategy for managing a stock portfolio. Download the S&P stock data from the course website, which contains price data for 1000 days of trading from 2001–2005 (in .csv or .mat format).
- (a) If an investor places all their money in the best single stock, what is the maximum possible earnings over the course of the 1000 trading days? What about if an equal amount of money is initially placed in each stock? Finally, what earnings would an investor achieve using a CRP strategy that allocates an equal percentage of the portfolio to each stock?
  - (b) Find a CRP strategy (i.e. a fixed allocation of funds) that performs better than the best single stock over the 1000 days. Include the strategy in a text file as a comma separated list containing the percentage of funds allocated to each stock. (Hint: Implement Gradient Descent. The following might be helpful for your projection step: <https://eng.ucmerced.edu/people/wwang5/papers/SimplexProj.pdf>.)

- (c) Of course, an investor would not be able to choose the best CRP strategy in hindsight. Implement online gradient descent to find a portfolio re-balancing strategy that an investor could actually use. What percentage gain were you able to achieve? Feel free to test multiple learning rates and report your best result. In practice, the learning rate itself would be learned based on prior data, or in an online way.

For parts (b) and (c) include a short discussion of your implementation. What are your gradient updates and how were they derived? Attach all code.

- §2 Consider an  $\alpha$ -strongly convex differentiable function  $f^1$  with the 2-norm of its gradient always bounded by  $G$ . Suppose we want to minimize  $f$  and let  $x^*$  denote its minimum. Show that the gradient descent algorithm with step size  $\frac{1}{\alpha(t+1)}$  satisfies

$$f\left(\frac{\sum_t x_t}{T}\right) - f(x^*) = \frac{G^2(1+\log T)}{2\alpha T}.$$

(Hint: Consider change in potential function  $\Phi(t) = \frac{t\alpha}{2}\|x_t - x^*\|^2$ , and use that  $\sum_{t \in \{1, \dots, T\}} \frac{1}{t} \leq 1 + \log T$ .)

- §3 Describe separation oracles for the following convex sets. Your oracles should run in linear time, assuming that the given oracles run in linear time (so you can make a constant number of black-box calls to the given oracles).

- (a) The  $\ell_1$  ball,  $\{x : \|x\|_1 \leq 1\}$ . Recall that  $\|x\|_1 = \sum_{i=1}^d |x_d|$ .
- (b) Any convex set  $A$  that we have a projection oracle for. I.e. we have an oracle to compute  $\arg \min_{x \in A} \|x - y\|_2$  for any  $y$ .
- (c) The  $\epsilon$ -neighborhood,  $E$  of any convex set  $A$ :

$$E = \{x : \exists y \in A \text{ with } \|x - y\|_2 \leq \epsilon\},$$

given a projection oracle for  $A$ .

- §4 Consider a set of  $n$  objects (images, songs, etc.) and suppose somebody has designed a *distance* function  $d(\cdot)$  among them where  $d(i, j)$  is the distance between objects  $i$  and  $j$ . We are trying to find a geometric realization of these distances. Of course, exact realization may be impossible and we are willing to tolerate a factor 2 approximation. We want  $n$  vectors  $u_1, u_2, \dots, u_n \in \mathbb{R}^n$  such that  $d(i, j) \leq \|u_i - u_j\|_2 \leq 2d(i, j)$  for all pairs  $i, j$ . Describe a polynomial-time algorithm that determines whether such  $u_i$ 's exist (and outputs them in the event that they do).

- §5 A  $k$ -sparse vector is any vector with at most  $k$  nonzero entries. Let  $\mathcal{S}_k$  be the set of all  $k$ -sparse vectors in  $\mathbb{R}^d$ . Show that, if  $\Pi$  is chosen to be a Johnson-Lindenstrauss embedding matrix (e.g. a scaled random Gaussian matrix) with  $s = O(\frac{k \log d}{\epsilon^2})$  rows then, with high probability,

$$(1 - \epsilon)\|\Pi x\|_2 \leq \|x\|_2 \leq (1 + \epsilon)\|\Pi x\|_2$$

---

<sup>1</sup>A differentiable function  $f$  is  $\alpha$ -strongly convex for  $\alpha > 0$  if for all  $x, y \in \mathbb{R}^n$  we have  $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\alpha}{2}\|y - x\|_2^2$ .

for all  $x \in \mathcal{S}_k$ , simultaneously.

**Hint:** You will want to use some result from the JL lecture as a black-box.

**Extra Credit:**

§1 (Extra Credit) Consider the following variant of the classical prophet inequalities where  $X_i$ s are no longer drawn independently: There is a known joint distribution over  $(X_1, \dots, X_n) \in \mathbb{R}_{\geq 0}^n$ . You are revealed the outcomes  $X_1, X_2, \dots$  one-by-one and on revelation of  $X_i$  you must immediately accept/reject it. Overall you can accept at most one of the  $n$  numbers. Prove that no algorithm can guarantee an expected accepted value better than  $\mathbb{E}[\max_i X_i]/n$  for every joint distribution.

§2 (Extra Credit) Attempt the puzzle at the [following url](#). If you solve it then first submit on codePost, and if the TAs approve it then you may also email Vincent Conitzer.