# COMPSCI 514 Spring 26: Final Review

## 1 Logistics

- The exam is closed book, no calculator or cheatsheets allowed.
- Held **in class Tuesday 5/12 from 1-3pm in CSL E110 (the lecture hall)**.
- Format will be similar to the midterms:
  - Q1. Set of TRUE/FALSE questions. Justification optional.
  - Q2. Set of numerical answer questions. Justification optional.
  - Q3-Q5. Longer 'homework style' questions.
  - Some extra credit worked into the questions. Or potentially Q5 will be all/mostly extra credit.

## 2 Studying Suggestions

- Focus primarily on practice questions. First from the past finals/midterms, then from the midterm review sheets and this sheet, questions I have asked in class, then from the textbooks if needed.
- Since we just had Midterm 2 on spectral methods, I would prioritize practice questions related to optimization and to the randomized algorithms unit.
- Make sure you understand all the homework and quiz problem solutions.
- Review slides/watch lectures as needed to fill in gaps in knowledge, but don't make this your primary means of studying.

## 3 Concepts to Study

**Probability and Randomized Algorithms**

- See Midterm 1 study guide.
- Do not need to memorize the analysis of specific algorithms like 2-level hashing, bloom filters, or count-min sketch. But should understand what problems these algorithms solve, what sorts of guarantees they give, and what general tools are used in the analysis.
- Should understand high level idea of LSH and how it is applied to nearest neighbor search, but won't be tested on s-curves or s-curve tuning.

**Spectral Methods**

- See Midterm 2 study guide.

**Optimization**

- Definition of gradient and connection to directional derivative.
- Ability to compute the gradient for very basic functions.
- Gradient descent.
- Convex function definition and corollary of what it implies about the gradient.
- Lipschitz function definition.
- Would not need to recreate the analysis of GD for convex Lipschitz functions and do not need to memorize the convergence theorem, but should understand the main ideas. Would be valuable to work through.
- Convex set definition, definition of projection, projected gradient descent for constrained optimization and why its analysis is essentially identical to that of gradient descent.
- Online optimization set up and online gradient descent (OGD).
- Regret definition. Why regret can be negative.
- Don't need to recreate OGD analysis or memorize the regret bound, but should understand the main ideas and how it compares to regular GD analysis.
- Idea of stochastic gradient descent (SGD) and why it is a special case of OGD. Don't need to know the analysis (we may not even get to it.)

# 4 Optimization Practice Questions

1. The difference of two convex functions $f(x)$ and $g(x)$ (i.e., $[f - g](x)$) is also convex. Always? Sometimes? Never?
2. The composition of two convex functions $f(x)$ and $g(x)$ (i.e., $[f \circ g](x)$) is also convex. Always? Sometimes? Never?
3. Let $\mathcal{S}$ be a convex set and let $f_{\mathcal{S}}(\vec{z}) = \begin{cases} 0 \text{ if } \vec{z} \in \mathcal{S} \\ 1 \text{ if } \vec{z} \notin \mathcal{S} \end{cases}$. Is $f_{\mathcal{S}}$ a convex function? Either prove that it is, or give a counterexample.
4. The sum of two $G$-Lipschitz functions is $2G$-Lipschitz. Always? Sometimes? Never?
5. The sum of two $G$-Lipschitz functions is $G$-Lipschitz. Always? Sometimes? Never?
6. Which of the following loss functions would our analysis of gradient descent for convex Lipschitz functions apply to? For each, explain why or why not.

$$f(\theta) = \frac{1}{\theta} + \theta \qquad g(\theta) = \sin(\theta) + \theta \qquad h(\theta) = 3 \cdot |\theta - 4|.$$

7. Let $\mathbf{X} \in \mathbb{R}^{n \times d}$ and $\vec{y} \in \mathbb{R}^n$ be fixed. Let $f(\vec{\theta}) = \|\mathbf{X}\vec{\theta} - \vec{y}\|_2^2$.

   (a) What is $\vec{\nabla} f(\vec{\theta})$?

   (b) What method other than gradient descent have we learned in class to minimize $f(\vec{\theta})$? What is the optimal solution $\vec{\theta}^*$?

8. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a $G$-Lipschitz function.

   (a) If $\theta^{(i+1)} = \theta^{(i)} - \eta \nabla f(\theta^{(i)})$, give an upper bound on $\|\theta^{(i+1)} - \theta^{(i)}\|_2$.

   (b) In our fixed step size gradient algorithm we set $t = \frac{R^2 G^2}{\epsilon^2}$ and $\eta = \frac{R}{G\sqrt{t}}$. Under these settings, what is the worst case increase in function value from step $i$ to step $i+1$.

   (c) Consider the case of projected gradient descent over a convex set $\mathcal{S}$. So $\theta^{(i+1)} = P_{\mathcal{S}}(\theta^{out})$ for $\theta^{out} = \theta^{(i)} - \eta \nabla f(\theta^{(i)})$. Show that the bound of (a) still holds.

9. Consider optimizing $f_1(x) = x^2$, $f_2(x) = (x-1)^2$ and $f_3(x) = (x+1)^2$ in an online fashion. What is $\theta^{off}$? What is the regret for the sequence of solutions $\theta^{(1)} = 0$, $\theta^{(2)} = .5$, $\theta^{(3)} = -.5$?

10. Consider a graph $G$ with Laplacian matrix $\mathbf{L}$. Consider the problem: $x_* = \arg\min_{x:\|x\|=1} x^T \mathbf{L} x$.

    (a) What is $x_*$? What value of $x_*^T \mathbf{L} x_*$ does it achieve?

    (b) Is the above optimization problem a convex optimization problem? Is it over a convex constraint set?

11. Let $\mathcal{S}$ be the set of all matrices $\mathbf{A}$ in $\mathbb{R}^{n \times n}$ with $\text{rank}(\mathbf{A}) \leq k$. Is $\mathcal{S}$ convex? Either prove that it is, or give a counterexample showing that it is not.

12. Consider some $\mathbf{v} \in \mathbb{R}^n$ and $\gamma \in \mathbb{R}$ and let $\mathcal{S}$ be the set of all matrices $\mathbf{A}$ in $\mathbb{R}^{n \times n}$ that have $\mathbf{v}$ as an eigenvector with eigenvalue $\gamma$. Is $\mathcal{S}$ convex? Either prove that it is, or give a counterexample showing that it is not.