CSCI 145 Problem 17: Random Rank or Best Rank?

August 27, 2026

Submission Instructions

Please upload your work to Gradescope by **11:59pm on DATE**.

- Turn in a single PDF of *handwritten* solutions. Typed (including LaTeX) solutions will not be accepted.
- Problems are grouped into three-week **cycles**. Once per cycle, you will meet with the grader and present the solution to *one* problem from that cycle, selected at random when you arrive; bring your work and be ready to explain any of them.

Problem 17: Random Rank or Best Rank? 2 points

LoRA replaces a weight update by $\mathbf{BA}$, where $\mathbf{B} \in \mathbb{R}^{m\times r}$ and $\mathbf{A} \in \mathbb{R}^{r\times k}$, with $1 \le r < k$. At initialization let $\mathbf{B}_0 = \mathbf{0}$ and let the rows of $\mathbf{A}_0$ be an orthonormal, uniformly random r-dimensional subspace of $\mathbb{R}^k$. Write $\mathbf{G} = \nabla_{\mathbf{W}}\mathcal{L}(\mathbf{W}_0)$ and consider one gradient step on $\mathbf{B}$ with $\mathbf{A}$ fixed.

A colleague claims that this first LoRA step is "essentially the best rank-r approximation to the full gradient step." Audit the claim. Derive the effective update to $\mathbf{W}$, compute the expected fraction of $\|\mathbf{G}\|_F^2$ retained by the random subspace, and compare its approximation error with the Eckart–Young–Mirsky optimum. Your verdict must address both an average random subspace and the worst possible alignment of a nonzero gradient whose rank is at most r.

State a defensible replacement for the colleague's claim and explain why the audit applies only to the initial step, before both factors have begun to adapt.

Results you may use. If $\mathbf{P} = \mathbf{A}_0^\top \mathbf{A}_0$, then $\mathbf{P}$ is an orthogonal rank-r projector and $\mathbb{E}[\mathbf{P}] = (r/k)\mathbf{I}$. You may quote Eckart–Young–Mirsky in Frobenius norm.