CSCI 145 Problem 15: The Other Norm, by Dimension Counting

July 25, 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 — so bring your work and be ready to explain any of them.

Problem 15: The Other Norm, by Dimension Counting

Two norms can disagree about what "close" means. The Frobenius norm $\|\mathbf{M}\|_F$ adds up the squared error in all nd entries, while the spectral norm

$$\|\mathbf{M}\|_2 = \max_{\mathbf{x}\neq \mathbf{0}} \frac{\|\mathbf{M}\mathbf{x}\|}{\|\mathbf{x}\|}$$

counts only the one direction $\mathbf{M}$ distorts the most. Compressing an image so it looks right on average and compressing it so no measurement from it is badly wrong are different goals, so they should have different answers. They have the same one: the truncated SVD $\mathbf{A}_k$ is optimal in both. Lecture proved the Frobenius case with projections and a trace; here we prove the spectral case by pigeonhole, with two subspaces too big to avoid each other.

Throughout, $\mathbf{A} \in \mathbb{R}^{n\times d}$ has SVD $\mathbf{A} = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^\top$ with $\sigma_1 \geq \cdots \geq \sigma_r > 0$ and orthonormal $\mathbf{u}_i \in \mathbb{R}^n$, $\mathbf{v}_i \in \mathbb{R}^d$; the rank-$k$ truncated SVD is $\mathbf{A}_k = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^\top$, for a budget $k < r$.

Part A: What the Truncation Achieves

Show that $\|\mathbf{A} - \mathbf{A}_k\|_2 = \sigma_{k+1}$ in two steps. First, evaluate $(\mathbf{A} - \mathbf{A}_k)\mathbf{v}_{k+1}$ using orthonormality of the $\mathbf{v}_i$ to see the norm is *at least* σ_{k+1}. Second, expand $\|(\mathbf{A} - \mathbf{A}_k)\mathbf{x}\|^2$ for any unit $\mathbf{x} \in \mathbb{R}^d$ to see it is *at most* σ_{k+1}. (Hint: since $\mathbf{A} - \mathbf{A}_k = \sum_{i=k+1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^\top$, only the components $\mathbf{v}_i^\top \mathbf{x}$ with $i > k$ survive, and their squares sum to at most 1.)

Part B: No Rank-k Matrix Does Better

Let $\mathbf{B} \in \mathbb{R}^{n\times d}$ be *any* matrix of rank at most k, and write

$$\ker(\mathbf{B}) = \{\mathbf{z} \in \mathbb{R}^d : \mathbf{B}\mathbf{z} = \mathbf{0}\}, \qquad V = \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{k+1}).$$

You may use two facts without proof: rank-nullity, which gives $\dim(\ker \mathbf{B}) \geq d - k$; and the pigeonhole principle for subspaces, which says two subspaces of $\mathbb{R}^d$ whose dimensions sum to more than d share a nonzero vector.

(i) Add up the two dimensions above and conclude that there is a *unit* vector $\mathbf{z} \in \ker(\mathbf{B}) \cap V$.

(ii) Write $\mathbf{z} = \sum_{i=1}^{k+1} c_i \mathbf{v}_i$ with $\sum_i c_i^2 = 1$, compute $\|\mathbf{A}\mathbf{z}\|^2$ in terms of the c_i and σ_i, and use $\sigma_i \geq \sigma_{k+1}$ for $i \leq k+1$ to show $\|\mathbf{A}\mathbf{z}\| \geq \sigma_{k+1}$.

(iii) Relate $(\mathbf{A} - \mathbf{B})\mathbf{z}$ to $\mathbf{A}\mathbf{z}$ and conclude $\|\mathbf{A} - \mathbf{B}\|_2 \geq \sigma_{k+1}$ for every such $\mathbf{B}$, which with Part A proves $\mathbf{A}_k$ is optimal in the spectral norm.

Part C: Why One Direction Was Enough

Part B produced exactly *one* bad direction $\mathbf{z}$, on which every rank-k competitor fails by at least σ_{k+1}. Explain why one direction settles the spectral norm, a maximum over directions, but cannot reproduce lecture's Frobenius bound $\sqrt{\sum_{i>k} \sigma_i^2}$, a sum over a whole orthonormal basis; what would pigeonhole have to produce instead, and does it? State in one sentence what this says about matching a proof technique to the norm you measure with.

Where this goes next. Two unrelated arguments, a trace identity and a subspace pigeonhole, single out the same $\mathbf{A}_k$ under two different notions of error, which is good evidence that truncating the SVD is the canonical way to spend a rank budget. The next two lectures truncate data (PCA) and weight *updates* (LoRA).