Deck 10: The randomized SVD

Math 7870: Topics in Randomized Numerical Linear Algebra

Spring 2026

Akil Narayan

The SVD

Recall: given a matrix $\boldsymbol{A} \in \mathbb{C}^{m \times n}$, with $m > n$, its (full) singular value decomposition is,

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^*, \qquad \boldsymbol{U} \in \mathbb{C}^{m \times m}, \qquad \boldsymbol{\Sigma} \in \mathbb{C}^{m \times n} \qquad \boldsymbol{V} \in \mathbb{C}^{n \times n}.$$

Both $\boldsymbol{U}$ and $\boldsymbol{V}$ are unitary, and $\boldsymbol{\Sigma}$ is diagonal with real, non-negative, and non-increasing entries.

The columns of $\boldsymbol{U}$ are the left singular vectors, the columns of $\boldsymbol{V}$ are the right singular vectors.

The diagonal elements of $\boldsymbol{\Sigma}$, $\sigma_1, \ldots, \sigma_n$, are the singular values of $\boldsymbol{\Sigma}$.

The singular values tell us a lot about compression/approximation of matrices. E.g.,

$$\boldsymbol{A}_r = \underset{\substack{\boldsymbol{B} \in \mathbb{C}^{m \times n} \\ \mathrm{rank}(\boldsymbol{B}) = r}}{\arg\min} \|\boldsymbol{A} - \boldsymbol{B}\|_2, \qquad \boldsymbol{A}_r = \boldsymbol{U}_r \boldsymbol{\Sigma}_r \boldsymbol{V}_r^*,$$

where $\boldsymbol{U}_r$ (resp. $\boldsymbol{V}_r$) is first r columns of $\boldsymbol{U}$, and $\boldsymbol{\Sigma}_r$ is the $r \times r$ principal submatrix of $\boldsymbol{\Sigma}$.

Computing the SVD

In the situation $m \gg n$, then computing the SVD requires $\mathcal{O}(mn^2)$ complexity.

But here's an observation. Let $\boldsymbol{Q} = \boldsymbol{U}_r$. Then the reduced SVD of $\boldsymbol{Q}^*\boldsymbol{A}$ is,

$$\boldsymbol{Q}\boldsymbol{U}_r = \boldsymbol{I}\boldsymbol{\Sigma}_r\boldsymbol{V}_r^*.$$

Therefore, here's an algorithm to compute a(n optimal) rank-r approximation to $\boldsymbol{A}$:

- Form $\boldsymbol{C} = \boldsymbol{Q}^*\boldsymbol{A} \in \mathbb{C}^{r \times n}$
- Compute the SVD of $\boldsymbol{C}$: $\boldsymbol{C} = \boldsymbol{W}\boldsymbol{\Lambda}\boldsymbol{Y}^*$.
- Let $\boldsymbol{U}_r = \boldsymbol{Q}\boldsymbol{W}$, $\boldsymbol{V}_r = \boldsymbol{Y}$.

Then $\boldsymbol{U}_r$ and $\boldsymbol{V}_r$ as computed above are exactly the first r truncated columns of $\boldsymbol{U}$ and $\boldsymbol{V}$, respectively, and $\boldsymbol{\Lambda}$ are the singular values of $\boldsymbol{A}$.

I.e., $\widetilde{\boldsymbol{A}}_r := \boldsymbol{Q}\boldsymbol{W}\Lambda\boldsymbol{Y}^*$ is a(n optimal) rank-r approximation to $\boldsymbol{A}$.

Step 1 above requires $\mathcal{O}(mnr)$ complexity, step 2 $\mathcal{O}(nr^2)$, and step 3 $\mathcal{O}(mr^2)$.

This is a signficant savings if $r \ll n$.

The problem: we don't know $\boldsymbol{Q}$.

The goal: a "rangefinder"

To more precisely identify what the goal should be, note that,

$$\|\boldsymbol{A} - \widetilde{\boldsymbol{A}}_r\|_2 = \|\boldsymbol{A} - \boldsymbol{Q}\boldsymbol{C}\|_2 = \|\boldsymbol{A} - \boldsymbol{Q}\boldsymbol{Q}^*\boldsymbol{A}\|_2.$$

Therefore, to approximately achieve Schmidt-Eckart-Young-Mirsky, it suffices to solve the problem for fixed r:

$$\text{Find } \boldsymbol{Q} \in \mathbb{C}^{\ell \times n} \text{ such that } \|\boldsymbol{A} - \boldsymbol{Q}\boldsymbol{Q}^*\boldsymbol{A}\|_2 \text{``} \approx \text{''} \sigma_{r+1}(\boldsymbol{A})$$

(for some $\ell \geqslant r$, with $\boldsymbol{Q}$ semi-unitary)
This is called the "rangefinder" problem.

The big reveal is that random matrices are excellent ways to compute rangefinders. E.g.,:

- Let $\boldsymbol{G} \in \mathbb{C}^{n \times \ell}$ be a random Gaussian matrix, with iid $\mathcal{N}(0, \frac{1}{\ell})$ entries.
- Compute $\boldsymbol{A}\boldsymbol{G} = \boldsymbol{Q}\boldsymbol{R}$. (The QR decomposition of $\boldsymbol{A}\boldsymbol{G}$.)
- On average, $\boldsymbol{Q}$ is a rank-r rangefinder with error $\approx (1 + \sqrt{r/(\ell - r)})\sigma_{r+1}(\boldsymbol{A}) + \frac{\sqrt{\ell}}{\ell - r}\|\boldsymbol{A} - \boldsymbol{A}_r\|_F$.

Here, "error" means the rangefinder error:

$$\|\boldsymbol{A} - \boldsymbol{Q}\boldsymbol{Q}^*\boldsymbol{A}\|_2 = \|(\boldsymbol{I} - \boldsymbol{P}_{\text{range}(\boldsymbol{A}\boldsymbol{G})})\boldsymbol{A}\|_2.$$

Proof idea

The proof is a nice exercise in linear algebra facts:

- The (rangefinder) error can be written through Schur complements:

$$\|(\boldsymbol{I} - \boldsymbol{P}_{\mathrm{range}(\boldsymbol{AG})})\boldsymbol{A}\|_2^2 = (\boldsymbol{A}^*\boldsymbol{A})/\boldsymbol{G},$$

  where $\boldsymbol{C}/\boldsymbol{D} = \boldsymbol{C} - (\boldsymbol{C}\boldsymbol{D})(\boldsymbol{D}^*\boldsymbol{C}\boldsymbol{D})^{-1}(\boldsymbol{C}\boldsymbol{D})^*$ is the Schur complement of $\boldsymbol{C}$ of $\boldsymbol{D}$.
- This Schur complement formulation yields the result,

$$\|(\boldsymbol{I} - \boldsymbol{P}_{\mathrm{range}(\boldsymbol{AG})})\boldsymbol{A}\|_2^2 \leqslant \|\boldsymbol{A} - \boldsymbol{A}_r\|_2^2 + \|\boldsymbol{\Sigma}_2\boldsymbol{V}_2^*\boldsymbol{G}(\boldsymbol{V}_1^*\boldsymbol{G})^\dagger\|_2^2,$$

  where $\boldsymbol{\Sigma}_2$ is the trailing $(n-r)\times(n-r)$ singular value matrix of $\boldsymbol{AG}$, and $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ are the first r and last $n-r$ columns, respectively of the right singular vector matrix of $\boldsymbol{A}$.
- Use randomized arguments to get concentraton around the mean.