Today

- Low-rank matrices in practice

- SVD $\Rightarrow kN^2$

- Range Finder

- RRQR

$\Rightarrow$ Interpolative Decomposition ID
2 Tools for Low-Rank Linear Algebra
Rephrasing Low-Rank Approximations


First, rephrase the LRA problem:

\[ A \approx QQ^T A \]
Using LRA bases

If we have an LRA basis $Q$, can we compute an SVD?

$$\beta = Q^\top A = \tilde{U} \Sigma V^\top$$

$$A \approx QQ^\top A = Q \tilde{U} \Sigma V^\top$$
Finding an LRA basis

How would we find an LRA basis?

\[ \mathbf{x} \sim N(0, I_n) \]

\[ A \in \mathbb{R}^{n \times n} \]

\[ A x = Q R \]

Below in \( (A^T A)^T \) is small, (but not that small) small values faster:

\[ (U \Sigma V^T)^T (U \Sigma V^T) = V \Sigma^2 V^T \]

\[ \in \text{ use this form to keep singular values small} \]
To avoid overflow: QR often enough.

$$A \approx$$

$$\exists \epsilon$$ With $T$ a QFT, the QFT will compute $\hat{\epsilon}$ in $O(n \log n)$ rather than $O(n^2)$. 
Giving up optimality

What problem should we actually solve then?
Recap: The Power Method

How did the power method work again?
How do we construct the LRA basis?

Put randomness to work:
Errors in Random Approximations

If we use the randomized range finder, how close do we get to the optimal answer?

For an $m \times n$ matrix and a target rank $k \geq 2$ and an oversampling parameter $p \geq 2$, and with $k \leq p \leq \min(m,n)$, with probability $1 - O(p^{-p})$,

$$\|A - QQ^*A\|_2 \leq \left(1 + 11\sqrt{k p \sqrt{\min(m,n)}}\right)\sigma_{\text{out}}.$$  

Halko, Martinsson, Tropp
A-posteriori and Adaptivity

The result on the previous slide was *a-priori*. Once we’re done, can we find out ‘how well it turned out’?
Rank-revealing/pivoted QR

Sometimes the SVD is too *good* (aka expensive)—we may need less accuracy/weaker promises, for a significant decrease in cost.

\[ A = QR \]

\[ A \Pi = Q \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix} \]

Interpolative decomposition
Interpolative Decomposition (ID)

Sometimes it would be helpful to know *which columns of A* contribute ‘the most’ to the rank.  
(rather than have the waters muddied by an orthogonal transformation like in QR)
What does the ID buy us?

Specifically: Name a property that the ID has that other factorizations do not have.
What does row selection mean for the LRA?
Leveraging the ID

Build a low-rank SVD with row extraction.