Today:

- ID demo
- Apply ID to the SVD
- Potential matrices
  - low rank
- Rank in terms of functions

$= A$

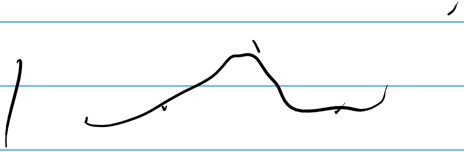
- HW due

$A = Q\underbrace{Q^T A}$

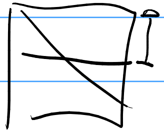
ID recap

A has rank k

$$A \approx A_{[:,J]}\, \underline{P}$$



Goal recap

$A\Pi = Q$ 

ID Q vs ID A

What does row selection mean for the LRA?

$$A \approx Q \overbrace{Q^T}^{} A \qquad Q$$

Run row ("transpose") ID on Q: $\quad Q \approx P\, Q_{[J,:]}$

$$A \approx P\, Q_{[J,:]}\, Q^T A$$

$$A_{[J,:]} \approx \underbrace{\cancel{P_{[J,:]}}}_{\text{ID}}\, Q_{[J,:]}\, Q^T A = Q_{[J,:]}\, Q^T A$$

$$P\, A_{[J,:]} = \underbrace{P\, Q_{[J,:]}}_{\approx Q}\, Q^T A \approx Q Q^T A \approx A$$

[Martinsson, Rokhlin, Tygert '06]

Demo: Interpolative Decomposition

- There is a slight tradeoff here--which?
- How do we use the ID in the context of low-rank range finding?

What does the ID buy us?

Name a property that the ID has over other factorizations.

All our randomized tools have two stages:

1. Find ONB of approximate range
2. Do actual work only on approximate range

Complexity?

What is the impact of the ID?

Ability to replace computing $Q^T A$ with $A_{[J,:]}$

SVD $A_{[j,:]} = U \Sigma V^T$

$$P A_{[j,:]} = \underbrace{P U}_{?} \Sigma V^T$$

not orth ..~

ID-based Complexity Reduction

$$\begin{array}{cc} & N \\ k & \boxed{A_{[J,:]}} \end{array}$$

How can we reduce factorization complexity with the ID?

- Assume we have P and J $A \approx P A_{[J,:]}$.

$$\underbrace{(A_{[J,:]})^T}_{N \times k} = \underbrace{\bar{Q}}_{N \times k} \underbrace{\bar{R}}_{k \times k}$$

$$\underbrace{Z}_{N \times k} = \underbrace{P}_{N \times k} \underbrace{R^T}_{k \times k}$$

$$Z = U \Sigma \bar{V}^T$$

$$N \times k \quad N \times k \quad k \times k \quad k \times k$$

Leveraging the ID for SVD (II)

In what way does this give us an SVD of A?

$$\underset{N\times k}{U}\ \underset{k\times k}{\Sigma}\ (\underset{N\times k}{\bar{Q}}\ \underset{k\times k}{\bar{V}})^T$$

$$= U\Sigma\bar{V}^T\bar{Q}^T$$

$$= Z\bar{Q}^T$$

$$= PR^T\bar{Q}^T$$

$$= PA_{[J,:]} \approx A$$

Leveraging the ID for SVD (III)

Q: Why did we need to do the row QR?

SVD $A_{[J,:]} = U\Sigma V^T$

$$P A_{[J,:]} = \underbrace{P\,U}_{?}\Sigma V^T$$

not orth ..

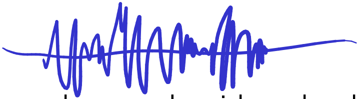
Where are we now?

- We have observed that we can make matvecs faster if the matrix has low-ish numerical rank
- In particular, it seems as though if a matrix has low rank, there is no end to the shenanigans we can play.
- We have observed that some matrices we are interested in (in some cases) have low numerical rank (cf. the point potential example)
- We have developed a toolset that lets us obtain LRAs and do useful work (using SVD as a proxy for "useful work") in $O(N \cdot K^{\alpha})$ time (assuming availability of a cheap matvec).

Next stop: Get some insight into *why* these matrices have low rank in the first place, to perhaps help improve our machinery even further.

Outline

Punchline



What do (numerical) rank and smoothness have to do with each other?

If all outputs of an op has a short expansion in some basis, then that is a continuous equivalent of "low rank".

poly, Fourier, eigenfunctions ...

Even shorter punchline?

Smoothing Operators

If the operations you are considering are *smoothing*, you can expect to get a lot of mileage out of low-rank machinery.

What types of operations are smoothing?

- d/dx : Derivatives

→ not smoothing

$$\mathcal{F}\{\partial_x f\}(\omega) = (i\omega)\,\mathcal{F}\{f\}(\omega)$$

- Integrals

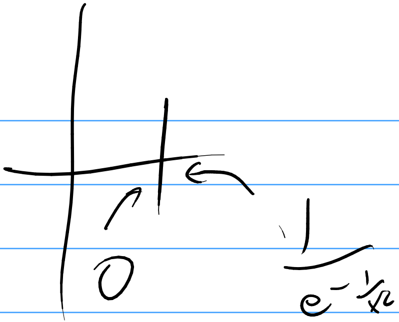
→ smoothing

$$\mathcal{F}\{\textstyle\int f\}(\omega) = (i\omega)^{-1}\,\mathcal{F}\{f\}(\omega)$$

Now: Consider some examples of smoothness, with justification.

How do we judge smoothness?

$C^\omega,\ C^\infty,\ \ldots\ C^k,\ \ldots\ C^0$ | by the remainder term in the Taylor series

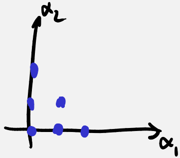

0
e^{-1/x}

Recap: Multivariate Taylor

$\mathbb{R}^n$

1D Taylor: $f(c+h) \approx \sum_{p=0}^{k} \frac{f^{(p)}(c)}{p!} h^p$

$$f(\vec{c}+\vec{h}) \approx \sum_{|\alpha| \le k} \frac{\partial^\alpha f(\vec{c})}{\alpha!} \vec{h}^\alpha$$



Multi-Index

$\alpha = (\alpha_1, \dots, \alpha_n)$

$\partial^\alpha f = \frac{\partial^{\alpha_1} \cdots \partial^{\alpha_n}}{\partial x_1^{\alpha_1} \, \partial x_n^{\alpha_n}} f$

$|\alpha| = |\alpha_1| + \dots |\alpha_n|$

$\alpha! = \alpha_1! \cdot \dots \alpha_n!$

$\vec{h}^\alpha = h_1^{\alpha_1} \dots h_n^{\alpha_n}$

Taylor and Error (I)

How can we estimate the error in a Taylor expansion?

(go back to 1D for simplicity)

$$\left| f(c+h) - \sum_{p=0}^{k} \frac{f^{(p)}(c)}{p!} h^p \right|$$

$$= \left| \sum_{p=k+1}^{\infty} \frac{f^{(p)}(c)}{p!} h^p \right| \leftarrow \text{If } f \text{ analytic}$$

Taylor and Error (II)

Now suppose that we had an estimate that $\left|\frac{f^{(p)}(c)}{p!}h^p\right| \leqslant \cancel{\ } \cdot \beta^p$

$$\left|\sum_{p=k+1}^{\infty} \frac{f^{(p)}(c)}{p!} h^p\right| \leq \sum_{p=k+1}^{\infty} \left|\frac{f^{(p)}(c)}{p!} h^p\right| = \sum_{p=k+1}^{\infty} \beta^p = \frac{1}{1-\beta} \cdot \beta^{k+1}$$

Useful for small enough h

β

Connect Taylor and Low Rank

Can Taylor help us establish low rank of an interaction?