Announcements

- Office hours moved to Fr 2/2 @ 10:30
- HW1

Goals

- wrap up $O(Nk^2)$ SVD
- why low rank?

Review

LRA:

$$A \approx \underbrace{\quad}_{\text{uncompressed}} \underbrace{\quad}_{\text{compressed form}}$$

		uncompressed	compressed form
proj form	→	Q	$Q^T A$
ID	→	P	$A_{(J,:)}$

↳ compressed form available cheaply for the ID

↳ "commutative property"

ID Q vs ID A

What does row selection mean for the LRA?

- Suppose we have PLRA: $A \approx Q Q^T A$ $\qquad Q \in \mathbb{R}^{N \times k}$
- Row ID: $Q \approx P Q_{(j)}$ $\qquad Q_{(j)} \in \mathbb{R}^{k \times k}$

$$A \approx \overbrace{P Q_{(j)}}^{Q} Q^T A$$

$$A_{(j)} \approx \underbrace{P_{(j)}}_{Id} Q_{(j)} Q^T A$$

$$P A_{(j)} \approx P Q_{(j)} Q^T A \approx A$$

[Martinsson, Rokhlin, Tygert '06]

ID: Remarks

Ω $N \times k$ Gaussian iid

$Y = A\Omega$ $N \times k$ $Y_{i+1} = AA^T Y_i$

$QR = Y$

$N \times k$

Slight tradeoff here: what?

More approximations.

How would we use the ID in the context of the range finder?

- "Compress" (i.e. throw away rows) as soon as possible
- Compute
- only decompress at the very end.

Demo: Interpolative Decomposition

What does the ID buy us?

Name a property that the ID has over other factorizations.

Commutes with other factorizations

All our randomized tools have two stages:

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

Complexity?

so far: $O(N^2 k)$

What is the impact of the ID?

Avoids formation of $Q^T A$

Leveraging the ID for SVD (I)

Build a low-rank SVD with row extraction.

- Get $\vec{j}$ and P so that $A \approx P A_{(\vec{j})}$
- $$\underbrace{\left(A_{\{\vec{j}\}}\right)^T}_{N \times k} = \underset{N\times k}{\bar{Q}} \; \underset{k\times k}{\bar{R}}$$
- Upsample $\bar{R}$: $$\underset{N\times k}{Z} = \underset{N\times k}{P} \; \underset{k\times k}{\bar{R}^T}$$
- SVD $$\underset{N\times k}{Z} = \underset{N\times k}{U} \; \underset{k\times k}{\Sigma} \; \underset{k\times k}{\bar{V}^T}$$

Leveraging the ID for SVD (II)

V: orth!

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

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

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

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

$$= P \tilde{R}^T \bar{Q}^T = P A_{(\tilde{J})} \approx A$$

Leveraging the ID for SVD (III)

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

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

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

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

Outline

Punchline

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

Smoothness: well-approximated by Taylor

If entire range of an operator has short-ish Taylor expansions (i.e. repr. wrt polynomials), that operator must be low rank

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?

∂ : no

$\int$: yes ←

Now: Consider some examples of smoothness, with justification.

How do we judge smoothness?

Taylor coeff.

Recap: Multivariate Taylor

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

multi-index: $\vec{p} = (p_1, \dots, p_n) \in \mathbb{N}_0^n$

$|\vec{p}| = p_1 + \dots + p_n$

$\vec{p}! = p_1! \cdot p_2! \cdots p_n!$

$\vec{x} \in \mathbb{R}^n$ $\quad \vec{x}^{\vec{p}} = x_1^{p_1} \cdots x_n^{p_n}$

$D^{\vec{p}} f = \frac{\partial^{|\vec{p}|} f}{\partial x_1^{p_1} \cdots \partial x_n^{p_n}}$

$f(\vec{c} + \vec{h}) \approx \sum_{|\vec{p}| \leq k} \frac{D^{\vec{p}} f(\vec{c})}{\vec{p}!} \vec{h}^{\vec{p}}$

Taylor and Error (I)

$$\sum_{p=k+1}^{\infty} \alpha^p = \frac{1}{1-\alpha}\alpha^{k+1}$$

$|\alpha|<1$

How can we estimate the error in a Taylor expansion?

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

if f analytic

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

(alternate reality: remainder terms, not here)

Taylor and Error (II)

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