Last time: Rand NLA : Low-rank approximation

<u>Goal</u> : Given $A \in \mathbb{R}^{m\times n}$, $k < \min\{m,n\}$ (assume $m \geq n$)

Find $$\min_{\text{rank}(A_k)\leq k} \|A - A_k\| \qquad (*)$$

Here we take $\|\cdot\|$ to be 2-norm or Frobenius norm

<u>Solution</u>:

Let SVD of A be $A = U\Sigma V^T = \sum_{i=1}^{n} \sigma_i u_i v_i^T$

where $U = [u_1 \; u_2 \; \cdots \; u_m] \in \mathbb{R}^{m\times m}$

$V = [v_1 \; v_2 \; \cdots \; v_n] \in \mathbb{R}^{n\times n}$

$\Sigma = \text{diag}(\sigma_1, \cdots, \sigma_n) \in \mathbb{R}^{m\times n}$

The optimal rank-k approximation of A is given by

$$A_k := U_k \Sigma_k V_k^T = \sum_{i=1}^{k} \sigma_i u_i v_i^T$$

where $U_k = [u_1 \; u_2 \; \cdots \; u_k] \in \mathbb{R}^{m\times k}$

$V_k = [v_1 \; v_2 \; \cdots \; v_k] \in \mathbb{R}^{n\times k}$

$\Sigma_k = \text{diag}(\sigma_1, \cdots, \sigma_k) \in \mathbb{R}^{k\times k}$

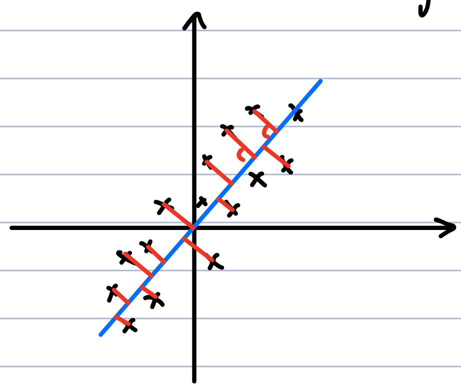
- An equivalent formulation of low-rank approximation is the Principal Component Analysis (PCA).

Let data points be m-dim vectors, stored in n columns of A. PCA aims to find k vectors whose span best contains the data points in A. We can assume these vectors are orthonormal basis and form the following problem:

$$\min_{\substack{Q\in\mathbb{R}^{m\times k}\\ Q^TQ=I_m}} \|A - \underbrace{QQ^T}_{=:P_Q} A\| \qquad (**)$$

projection onto span of Q

that is, $A \approx QQ^TA$, Q is an approximate basis for the range of A



PCA is essentially low-rank approximation:

- On one hand. $\displaystyle\min_{\operatorname{rank}(\hat{A})\le k} \|A-\hat{A}\| \le \min_{\substack{Q\in\mathbb{R}^{m\times k}\\ Q^TQ=I_m}} \|A - \underbrace{QQ^TA}_{\operatorname{rank}(QQ^TA)\le k}\|$

- On the other hand, we take $Q = [u_1 \cdots u_k] \in \mathbb{R}^{m\times k}$

then $$\underbrace{QQ^TA}_{=:C} = [u_1 \cdots u_k]\begin{bmatrix} -u_1^T- \\ \vdots \\ -u_k^T- \end{bmatrix}[u_1 \cdots u_m]\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_n\end{bmatrix}\begin{bmatrix} -v_1^T- \\ \vdots \\ -v_n^T- \end{bmatrix}$$

$$= [u_1 \cdots u_k]\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_k\end{bmatrix}\begin{bmatrix} -v_1^T- \\ \vdots \\ -v_k^T- \end{bmatrix} \longleftarrow \text{truncated SVD of } A$$

The optimal $Q_k = U_k = [u_1 \cdots u_k]$

- How to efficiently compute low-rank approximation of A?
  - Apply SVD to A then truncate $\longleftarrow$ expensive! $O(mn^2)$
  - Workaround: two stage algorithm from PCA

<u>Stage A</u>: compute orthonormal basis Q whose span approximates $\operatorname{span}(A)$

$$\Rightarrow \quad A \approx QQ^TA \quad \text{and} \quad Q^TQ = I$$

<u>Stage B</u>: Compute $C = Q^TA \in \mathbb{R}^{k\times n}$ ← $O(kmn)$

then compute SVD of C (use whatever method):

$$B = \tilde{U}\Sigma V^T \quad \leftarrow O(k^2n)$$

then $A \approx (Q\tilde{U})\Sigma V^T$ ← $O(k^2m)$

- How to find Q? (Randomized "sketch")

<u>Idea</u>: Approximate $\operatorname{span}\{u_1, \cdots, u_k\}$. the top k singular vectors of A, with a single power iteration

<u>Intuition 1</u>: Suppose A has exactly rank k, so the best rank k approximation of A is A itself. Suppose $\sigma_k > 0$

Compute $Y = A\Omega$, $\Omega \in \mathbb{R}^{n\times k}$

$$[Y_1 \cdots Y_k] = [u_1 \cdots u_k]\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_k\end{bmatrix}\begin{bmatrix} - v_1^T - \\ \vdots \\ - v_k^T - \end{bmatrix}[\omega_1 \cdots \omega_k]$$

$\omega_i \in \mathbb{R}^n$ random vectors

Y has linearly independent columns as long as $\operatorname{rank}(V^T\Omega) = k$

For i.i.d. normal entries of Ω, this happens almost surely.

$V^T\Omega$ still has iid normal entries

To get orthonormal basis, we compute $Y = QR$

Clearly $\|A - QQ^TA\|_2 = 0 = \sigma_{k+1}$ a.s.

Intuition 2: Now let $A = \underbrace{\sum_{i=1}^{k} \sigma_i u_i v_i^T}_{=: \hat{A}_k} + \underbrace{\sum_{j=k+1}^{n} \sigma_j u_j v_j^T}_{=: E}$

$(\|E\|_2 = \sigma_{k+1} \text{ small})$

then $$y_i = A w_i = \overbrace{\hat{A}_k w_i}^{=: \tilde{y}_i} + E w_i$$

$$Y = \tilde{Y} + E\Omega$$

basis for span$(\hat{A}_k)$

small perturbation

The span of Y is a good approximation to $\tilde{Y}$ with high probability as long as we oversample, i.e. take $\underline{\Omega \in \mathbb{R}^{n \times (k+p)}}$ instead (p is oversampling parameter)

Algorithm: (Randomized SVD)

Stage A: 1) Generate iid Gaussian random matrix

Optimal range finder

$\Omega \in \mathbb{R}^{n \times (k+p)}$

2) Compute $Y = A\Omega \in \mathbb{R}^{m \times (k+p)}$

3) Compute QR fact. $Y = \hat{Q}_{k+p} R$, $\hat{Q}_{k+p} \in \mathbb{R}^{m \times (k+p)}$

Stage B: 1) Compute $B = \hat{Q}_{k+p}^T A \in \mathbb{R}^{(k+p) \times n}$

Post-processing (truncated SVD)

2) Compute SVD: $B = \tilde{U}_{k+p} \hat{\Sigma}_{k+p} \hat{V}_{k+p}^T$

3) Compute $\hat{U}_{k+p} = \hat{Q}_{k+p} \tilde{U}_{k+p} \in \mathbb{R}^{m \times (k+p)}$

Output: $\hat{A}_k := \hat{U}_k \hat{\Sigma}_k \hat{V}_k^T = \sum_{i=1}^{k} \hat{\sigma}_i \hat{u}_i \hat{v}_i^T$

Error analysis: Let $P_Y := \hat{Q}_{k+p} \hat{Q}_{k+p}^T$

We split the final error into two parts

$$\| A - \hat{A}_k \| \le \underbrace{\| A - P_Y A \|}_{\substack{\text{approximation err} \\ \text{of ranger finder} \\ \text{(Stage A)}}} + \underbrace{\| P_Y A - \hat{A}_k \|}_{\substack{\text{truncation err} \\ \text{(stage B)}}}$$

Given $\hat{Q}_{k+p}$, Stage B is computing the best rank k approximation of $P_Y A := \hat{Q}_{k+p} \hat{Q}_{k+p}^T A$

(b.c. $P_Y A = \hat{Q}_{k+p} \tilde{U}_{k+p} \hat{\Sigma}_{k+p} \hat{V}_{k+p}^T \longleftarrow$ SVD of $P_Y A$

and the output truncates to the top k modes)

Thus $\| P_Y A - \hat{A}_k \|$ is easy to controll:

optimal rank k approx. of A

$$\| P_Y A - \hat{A}_k \| \le \| P_Y (A - A_k) \| \le \| A - A_k \| \le \begin{cases} \sigma_{k+1}, & \|\cdot\| = \|\cdot\|_2 \\ \left(\sum_{j=k+1}^{n} \sigma_j^2 \right)^{1/2}, & \|\cdot\| = \|\cdot\|_F \end{cases}$$

$\hat{A}_k$ is optimal

projection is a contraction (length shrinking)

It remains to analyze the error in Stage A, i.e. $\| A - P_Y A \|$

- Special case: $\text{rank}(A) = k$, $\Omega \in \mathbb{R}^{n \times k}$

  Then $Y = A\Omega = U_k \Sigma_k \underbrace{(V_k^T \Omega)}$

  $\text{rank}(V_k^T \Omega) = k$ almost surely

  so $\text{range}(Y) = \text{range}(U_k) = \text{range}(A)$ almost surely

  $\Rightarrow \| A - P_Y A \| = 0$

- General case: $\text{rank}(A) > k$. oversampling $p > 1$

$$Y = A\Omega = [U_k, \bar{U}] \begin{bmatrix} \Sigma_k & \\ & \bar{\Sigma} \end{bmatrix} \begin{bmatrix} V_k^T \\ \bar{V}^T \end{bmatrix} \Omega$$

$$= \underbrace{U_k \Sigma_k (V_k^T \Omega)}_{=: Y_k} + \bar{U}\underbrace{\bar{\Sigma}(\bar{V}^T \Omega)}_{\text{perturbation } O(\sigma_{k+1})}$$

$\text{range}(Y_k) = \text{range}(U_k)$ as long as $V_k^T \Omega$ has full row rank

Thm (Perturbation lemma)

For $\|\cdot\| = \|\cdot\|_2$ or $\|\cdot\|_F$,

and any $\Omega \in \mathbb{R}^{n\times(k+p)}$, s.t. $V_k^T \Omega \in \mathbb{R}^{k\times(k+p)}$ has full row rank, then

$$\|(I - P_Y)A\|^2 \le \|\bar{\Sigma}\|^2 + \|\bar{\Sigma}\underbrace{(\bar{V}^T\Omega)}_{\Omega_2 :=}\underbrace{(V_k^T \Omega)}_{\Omega_1 :=}{}^{\dagger}\|^2$$

$\mathbb{R}^{(n-k)\times(k+p)}$ pseudoinverse

Ω_1, Ω_2 iid Gaussian entry

Use this lemma, we can prove average error bound:

$$\mathbb{E}\left[\|(I-P_Y)A\|_F^2\right] \le \|\bar{\Sigma}\|_F^2 + \mathbb{E}\left[\|\bar{\Sigma}\,\Omega_2\,\Omega_1^{\dagger}\|_F^2\right]$$

Conditioned on Ω_1, we have

$$\mathbb{E}[\|\bar{\Sigma}_2 \Omega_2 \Omega_1^{\dagger}\|_F^2 \mid \Omega_1] = \|\bar{\Sigma}\|_F^2 \|\Omega_1^{\dagger}\|_F^2$$

$\mathbb{E}[\|SGT\|_F^2] = \|S\|_F^2 \|T\|_F^2$ (exercise)

Hence,

$$\mathbb{E}[\|(I-P_Y)A\|_F^2] \le \|\bar{\Sigma}\|_F^2 + \|\bar{\Sigma}_F\|_F^2 \, \mathbb{E}\left[\|\Omega_1^{\dagger}\|_F^2\right]$$

Use Inverse Wishart distribution

$$= \left(1 + \frac{k}{p-1}\right)\|\bar{\Sigma}\|_F^2$$

$$= \left(1 + \frac{k}{p-1}\right)\left(\sum_{j>k} \sigma_j^2\right)$$

$\|\Omega_1^{\dagger}\|_F^2 = \text{tr}[(\Omega_1^T \Omega_1)^{-1}]$

Remark: Randomness ensure $\mathbb{E}[\|\Omega_1^{\dagger}\|_F^2]$ is small

Similar bound can be obtained for 2-norm

Thm $A \in \mathbb{R}^{m\times n}$. $k \ge 2$, $p \ge 2$ s.t. $k+p \le \frac{1}{2}\min\{m,n\}$

Then Stage A produces $m\times(k+p)$ orthonormal basis, s.t.

$$\mathbb{E}\left[\|A - \hat{Q}_{k+p}\hat{Q}_{k+p}^T A\|_2\right] \le \sigma_{k+1}\left[1 + \frac{4\sqrt{k+p}}{p-1}\sqrt{\min\{m,n\}}\right]$$

If $4 \log 4 \leq p \leq \min\{m, n\}$, then

$$\|A - \hat{Q}_{k+p} \hat{Q}_{k+p}^T A\|_2 \leq 6_{k+1} [1 + 9\sqrt{k+p}\sqrt{\min\{m,n\}}]$$

with probability $\geq 1 - 3p^{-p}$ ☒

For $p = 10$. failure probability $\leq 3 \times 10^{-10}$!

- In the proof above, essentially we want

$$\|(\bar{V}^T \Omega)(V_k^T \Omega)^+\| = O(1)$$

In the special case $k+p = n$, and $\Omega = I_n$

$$\|(\bar{V}^T \Omega)(V_k^T \Omega)^+\| = \|\bar{V}^T V_k\| = 0$$

orthogonal basis

In general, we have

$$\bar{V}^T \Omega = \begin{bmatrix} - \bar{v}_1^T - \\ \vdots \\ - \bar{v}_{n-k}^T - \end{bmatrix} \Omega = \begin{bmatrix} \bar{v}_1^T \Omega \\ \vdots \\ \bar{v}_{n-k}^T \Omega \end{bmatrix} \in \mathbb{R}^{(n-k) \times (k+p)}$$

$$V_k^T \Omega = \begin{bmatrix} v_1^T \Omega \\ \vdots \\ v_k^T \Omega \end{bmatrix} \in \mathbb{R}^{k \times (k+p)}$$

In order for $\|\bar{V}^T \Omega (V_k^T \Omega)^+\| \leq \|\bar{V}^T \Omega\| \, \|(V_k^T \Omega)^+\| = O(1)$

we need $\|\bar{v}_i^T (\frac{1}{\sqrt{n}} \Omega)\| \approx \|\bar{v}_i\|$, $\|v_j^T (\frac{1}{\sqrt{n}} \Omega)\| \approx \|v_j\|$, $\forall i, j$

- In other words, we can view $\frac{1}{\sqrt{n}} \Omega$ as a map from $\mathbb{R}^n$ to $\mathbb{R}^{k+p}$, and hope that the norm (or angle) of $v_1, \cdots, v_k, \bar{v}_1, \cdots, \bar{v}_{n-k} \in \mathbb{R}^n$ are approximately preserved.

We want the choice of Ω is such that it "sketch" every given vector in an low-dimensional subspace with small distortion, without knowing the subspace in advance. Such choice of Ω is called <u>oblivious subspace embedding</u>.

Intuitively, we know $\|v^T(\frac{1}{\sqrt{n}}\Omega)\|_2 \approx \|v^T\|_2$ is requiring $\frac{1}{n}\Omega\Omega^T \approx I_k$, which can be seen from

$$\frac{1}{n}\mathbb{E}[(\Omega\Omega^T)_{ij}] = \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}[w_{ik}w_{jk}] = \delta_{ij}$$

- Other than Gaussian matrices, there are other options such as subsampled randomized Hadamard (or Fourier) transforms, which allow fast evaluation of $A\Omega$.

i.e. $\Omega = \sqrt{\frac{n}{l}}\,DFR \quad \in \mathbb{C}^{n\times l}$ (assume complex A for simplicity)

D: $n\times n$ diagonal iid uniform on unit circle

F: $n\times l$ unitary discrete FFT transform

R: $n\times l$ each column sample w/o replacement from I_n

when $l \sim k\log(k)$, we have

$$0.4 \le \sigma_k(V^*\Omega),\ \sigma_1(V^*\Omega) \le 1.48$$