

Homework 3

Due Date: 04-03-2024

Assignments are to be submitted through Canvas, and should be individual work. You can discuss the problems, but should submit individually. Preferably typewritten.

Problem 1. Power method analysis

Let \mathbf{A} be an $n \times d$ matrix and \mathbf{x} a unit length vector in \mathbb{R}^d with $|\mathbf{x}^\top \mathbf{v}_1| \geq \eta$, where $\eta > 0$ and \mathbf{v}_1 is the top right singular vector of \mathbf{A} . Let W be the space spanned by the right singular vectors of \mathbf{A} corresponding to singular values greater than $(1 - \epsilon)\sigma_1$. Let \mathbf{z} be the unit vector after $q = \frac{\log(1/\epsilon\eta)}{2\epsilon}$ iterations of the power method, namely

$$\mathbf{z} = \frac{(\mathbf{A}^\top \mathbf{A})^q \mathbf{x}}{\|(\mathbf{A}^\top \mathbf{A})^q \mathbf{x}\|}$$

Then, show that \mathbf{z} has a component of at most ϵ perpendicular to W .

(Note: if \mathbf{x} is a Gaussian vector, we saw in Lecture 12 that $\eta \approx 1/d^3$.)

Hints: (i) Consider writing \mathbf{z} as a linear combination of the right singular vectors \mathbf{v}_i 's. (see slides 14 and 15 in Lecture 12).

(ii) Let $\sigma_1, \dots, \sigma_m$ be the singular values of \mathbf{A} that are $\geq (1 - \epsilon)\sigma_1$ for some m .

(iii) Use hints (i) and (ii) to write out the component of \mathbf{z} that is perpendicular to W . Find an upper bound to its squared length.

(iv) Use the first inequality in slide 16 of Lecture 12, and the value of q above to show that this length is at most ϵ .

Problem 2. Hutchinson's estimator analysis

The *Hanson-Wright inequality* is defined as: Given a symmetric matrix $\mathbf{B} \in \mathbb{R}^{n \times n}$ and random vector $\mathbf{z} \in \mathbb{R}^n$ with mean zero, i.i.d sub-Gaussian entries, and constant sub-Gaussian parameter C , we have for $t \geq 0$:

$$\Pr \left(\left| \mathbf{z}^\top \mathbf{B} \mathbf{z} - \mathbb{E}[\mathbf{z}^\top \mathbf{B} \mathbf{z}] \right| \geq t \right) \leq 2 \exp \left(-c \cdot \min \left(\frac{t^2}{\|\mathbf{B}\|_F^2}, \frac{t}{\|\mathbf{B}\|_2} \right) \right),$$

for some universal constant $c > 0$ that only depending on C .

Using this, show that the Hutchinson's estimator, $\tilde{\text{Tr}}_m(\mathbf{A}) = \frac{1}{m} \sum_{l=1}^m \mathbf{x}_l^\top \mathbf{A} \mathbf{x}_l$ where $\mathbf{A} \in \mathbb{R}^{d \times d}$ is SPD, and $\mathbf{x}_l, l = 1, \dots, m$ are random vectors with mean zero, i.i.d sub-Gaussian entries, if $m \geq \frac{c \log(2/\delta)}{\epsilon^2}$, then

$$\Pr \left[\left| \tilde{\text{Tr}}_m(\mathbf{A}) - \text{Tr}(\mathbf{A}) \right| \leq \epsilon \mid \text{Tr}(\mathbf{A}) \right] \geq 1 - \delta.$$

Hint: Consider applying the Hanson-Wright inequality to a block-diagonal matrix with repeated diagonal entries. We discussed this in the class.

Problem 3. Tensors

Consider the following tensor:

$$\mathcal{A}_{:, :, 1} = \begin{bmatrix} 3 & 9 & 1 \\ 8 & 2 & 1 \\ 4 & 3 & 9 \end{bmatrix} \text{ and } \mathcal{A}_{:, :, 2} = \begin{bmatrix} 6 & 9 & 5 \\ 5 & 6 & 4 \\ 1 & 4 & 1 \end{bmatrix}$$

- (a) Find $\mathcal{A}_{2, :, :}$ and $\mathcal{A}_{2, 3, :}$
- (b) Write $\text{vec}(\mathcal{A})$
- (c) Write $\mathbf{A}_{(2)}$ and $\mathbf{A}_{(3)}$
- (d) Compute $\|\mathcal{A}\|_F^2$.

Problem 4. Khathri-Rao product properties:

Given the Kronecker product properties:

$$\begin{aligned} (\mathbf{B} \otimes \mathbf{A})^\top &= \mathbf{B}^\top \otimes \mathbf{A}^\top \\ (\mathbf{B} \otimes \mathbf{A})(\mathbf{D} \otimes \mathbf{C}) &= (\mathbf{B}\mathbf{D}) \otimes (\mathbf{A}\mathbf{C}) \end{aligned}$$

Prove:

- $(\mathbf{B} \odot \mathbf{A})^\top (\mathbf{B} \odot \mathbf{A}) = \mathbf{B}^\top \mathbf{B} * \mathbf{A}^\top \mathbf{A}$
- $(\mathbf{B} \otimes \mathbf{A})(\mathbf{D} \odot \mathbf{C}) = (\mathbf{B}\mathbf{D}) \odot (\mathbf{A}\mathbf{C})$

Note that ‘*’ is the elementwise (Hadamard) product.

Problem 5. CP-ALS and randomized CP

Download the Uber tensor from (https://gitlab.com/tensors/tensor_data_uber). This is a 4-way tensor of size $183 \times 24 \times 1,140 \times 1,717$.

- (a) Run and time CP ALS for ranks 5:5:20. Plot the relative errors.

(You can use `cp_als` function from tensor toolbox or `parafac` function from tensorly package)

- (b) Run and time CP-ARLS-Mix for the same set of ranks. Plot the relative errors. How do these compare to CP ALS?

(You can use `cp_arls` function with ‘mix’ parameter set to true from tensor toolbox or `randomized_parafac` function from tensorly package)