CSE 392: Matrix and Tensor Algorithms for Data Spring 2024

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

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

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

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

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

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

(iii) Use hints (i) and (ii) to write out the component of $\boldsymbol{z}$ 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 this at most ϵ.

Problem 2. Hutchinson's estimator analysis

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

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

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

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

$$\Pr\left[\left|\tilde{\mathrm{Tr}}_m(\boldsymbol{A}) - \mathrm{Tr}(\boldsymbol{A})\right| \leq \epsilon|\mathrm{Tr}(\boldsymbol{A})|\right] \geq 1 - \eta.$$

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 $\mathrm{vec}(\mathcal{A})$

(c) Write $\boldsymbol{A}_{(2)}$ and $\boldsymbol{A}_{(3)}$

(d) Compute $\|\mathcal{A}\|_F^2$.

Problem 4. Khathri-Rao product properties:

Given the Kronecker product properties:

$$(\boldsymbol{B} \otimes \boldsymbol{A})^\top = \boldsymbol{B}^\top \otimes \boldsymbol{A}^\top$$
$$(\boldsymbol{B} \otimes \boldsymbol{A})(\boldsymbol{D} \otimes \boldsymbol{C}) = (\boldsymbol{B}\boldsymbol{D}) \otimes (\boldsymbol{A}\boldsymbol{C})$$

Prove:

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

Note that '$*$' is the elementwise (Hadamard) product.

Problem 5. CP-ALS and randomized CP

Download the Monkey BMI data from (`https://gitlab.com/tensors/tensor_data_monkey_bmi`).
The `data.mat` contains a 3-way tensor of size $43 \times 200 \times 88$.

(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)

In Matlab, you will have to apply `tensor` function to convert matlab array to a tensor object.

You can use the '`viz_monkey_bmi_cp`' function to visualize the CP factors.
(`viz_monkey_bmi_cp(M, angle)`, where M is CP output tensor and angle is an array inside data.mat file)