Exercise 7 – Low-rank approximation techniques

Prof. Dr. D. Kressner
A. Cortinovis

Problem 1: Column sampling error in the Frobenius norm

The goal of this exercise is proving that the lemma on column sampling from the slides has the following form for the Frobenius norm:

$$\|A - QQ^T A\|_F^2 \leq \|A - \mathcal{T}_r(A)\|_F^2 + 2\sqrt{r}\|AA^T - CC^T\|_F.$$

(a) Prove that

$$\|A - QQ^T A\|_F^2 \leq \|A - \mathcal{T}_r(A)\|_F^2 + \Big| \|\mathcal{T}_r(A)\|_F^2 - \|C^T Q\|_F^2 \Big| + \Big| \|A^T Q\|_F^2 - \|C^T Q\|_F^2 \Big|.$$

(b) Prove that $\Big| \|A^T Q\|_F^2 - \|C^T Q\|_F^2 \Big| \leq \sqrt{r}\|AA^T - CC^T\|_F$.

(c) Prove that $\Big| \|\mathcal{T}_r(A)\|_F^2 - \|C^T Q\|_F^2 \Big| \leq \sqrt{r}\|AA^T - CC^T\|_F$;

Hint: The Ky Fan eigenvalue inequality

$\lambda_1(X+Y) + \ldots + \lambda_k(X+Y) \leq \lambda_1(X) + \ldots + \lambda_k(X) + \lambda_1(Y) + \ldots + \lambda_k(Y)$

which holds for $n \times n$ matrices X, Y and $k = 1, \ldots, n$ could be useful.

Problem 2: Implementation of column sampling strategies

Implement both sampling strategies discussed in the lecture in Matlab:

(a) sampling based on column norms;

(b) sampling based on $\|V_k(\ell, :)\|_2$.

Find an example for which the second strategy performs much better than the first.

Hint: Make the norms $\|V_k(\ell, :)\|_2$ very different while keeping the column norms balanced.

Problem 3: Hoeffding's inequality and trace estimation

(a) Deduce Hoeffding's inequality from the McDiarmid's inequality that you have seen in the lecture: For $X_1, \ldots, X_n$ independent random variables such that $a \leq X_i \leq b$ for $i = 1, \ldots, n$, we have

$$\mathbb{P}\left(\left| \frac{1}{n}\sum_{i=1}^n X_i - \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n X_i \right] \right| \geq \varepsilon \right) \leq 2\exp\left(-\frac{2n\varepsilon^2}{(b-a)^2} \right).$$

(b) Let Y be a Rademacher random vector of length n, that is, its entries are ± 1 with probability $\frac{1}{2}$ independently from one another. Prove that, for a symmetric matrix $A \in \mathbb{R}^{n \times n}$, we have $\mathbb{E}[Y^T AY] = \text{trace}(A)$.

(c) Point (b) suggests that we can approximate trace(A) by taking N Rademacher sample vectors $Y_1, \ldots, Y_N$ and computing $\text{trace}_N(A) := \frac{1}{N}\sum_{i=1}^N Y_i^T AY_i$. What value of $N = N(\varepsilon, \delta, A)$ ensures that

$$\mathbb{P}(|\text{trace}(A) - \text{trace}_N(A)| \geq \varepsilon) \leq \delta?$$

Hint: Apply Hoeffding's inequality to the random variables $X_i := Y_i^T AY_i$.