Przemysław Uznański 02/12/2019

Algorithms for Big Data

Fall Semester 2019

Exercise Set 8

Consider a regression problem of

$$\arg\min_{X} \quad \|AX - B\|_F \tag{1}$$

where $A \in \mathbb{R}^{n\times d}$, $B \in \mathbb{R}^{n\times m}$ and $X \in \mathbb{R}^{d\times m}$.

Exercise 1:
Show that $X = A^\dagger B$ is a solution to (1).

Exercise 2:
Show that X from previous exercise minimizes $\|X\|_F$ among all the solutions.

We move to low-rank approximation:

$$A_k = \arg\min_{B:\mathrm{rank}(B)\le k} \|A - B\|_F$$

Exercise 3:
Show that Σ_k (as defined on the lecture) is a low-rank approximation to Σ wrt to Frobenius norm (that is it solves the problem for diagonal matrices).

Exercise 4:
Use previous exercise to show that $A_k = U\Sigma_k V^T$ is indeed low-rank approximation to $A = U\Sigma V^T$.

We move to Fourier transform. Let $\omega = e^{-\frac{2\pi}{n}}$. Let F be such that $F_{ij} = \frac{1}{\sqrt{n}}\omega^{ij}$. Then $\hat{a} = Fa$ is a (Discrete) Fourier transform of a.

Exercise 5:
Show how to compute $\hat{a}$ in time $\mathcal{O}(n \log n)$ (you can assume n is a power of two).

Exercise 6:
Show that $\|a\|_2 = \|\hat{a}\|_2$.

Exercise 7:
Let $\hat{a}_k$ is $\hat{a}$ with all but k largest-magnitude coefficients zeroed. Show that $a_k = F^{-1}\hat{a}_k$ is a solution to

$$\arg\min_{x:fs(x)\le k} \|a - x\|_2$$

where $fs(x) = \|\hat{x}\|_0$ is the size of Fourier support.