0368-3248-01-Algorithms in Data Mining **Fall 2012**

Lecture 10: Assignment 3

Lecturer: Edo Liberty

Warning: *This note may contain typos and other inaccuracies which are usually discussed during class. Please do not cite this note as a reliable source. If you find mistakes, please inform me.*

---

1 Randomized meta-algorithms

setup

In this question we assume the common case where we have an input $x \in X$ and we wish to approximate a function $f : X \to \mathbb{R}^+$ (i.e. $\forall x \;\; f(x) \geq 0$). For that we have a black box randomized algorithm $A : X \to \mathbb{R}^+$ such that $\mathbb{E}[A(x)] = f(x)$. The questions ask you to designing meta algorithms using A as a black box.

question

1. Show that
$$\Pr[A(x) \geq 3f(x)] \leq \frac{1}{3}$$

2. Assume that for all x we have that $\mathrm{Var}[A(x)] \leq c \cdot [f(x)]^2$. Describe an algorithm B_2 such that for any two constants $\varepsilon, \delta > 0$:
$$\Pr[|B_2(x) - f(x)| \geq \varepsilon f(x)] \leq \delta$$

3. Assume that $\Pr[|A(x) - f(x)| \leq t|] \geq \frac{1}{2} + \eta$ for some fixed value $\eta > 0$. In words, the algorithm gets an additive approximation t with probability slightly better than 1/2. (Here we do not assume anything on the variance of $A(x)$). Design and algorithm B_3 such that for any prescribed $\delta > 0$
$$\Pr[|B_3(x) - f(x)| \leq t|] \geq 1 - \delta$$
That means the algorithm achieves the same additive approximation with probability arbitrary close to one.

2 SVD and the power method

setup

Here we will prove some basic facts about singular values, matrices, and the power method. For the reminder of the question we assume $A \in \mathbb{R}^{m \times n}$ is an arbitrary matrix. For convenience and w.l.o.g. assume $m \leq n$. Also, denote by $\sigma_1 \geq \ldots \sigma_m \geq 0$ the singular values of A.

question

1. Let $P \in \mathbb{R}^{m \times m}$ and $Q \in \mathbb{R}^{n \times n}$ be unitary matrices. Show that $\|PAQ\|_{fro} = \|A\|_{fro}$. Hint, begin with the case where one of the matrices P or Q are the identity matrix.

2. Using the above show that for any matrix A we have that

$$\|A\|_{fro} = \sqrt{\sum_{i=1}^{m} \sigma_i^2}.$$

   It might help you to show that $\|A\|_{fro}^2 = tr(AA^T)$ where $tr(\cdot)$ stands for the matrix trace.

3. The numerical rank of a matrix $\rho(A) = \frac{\|A\|_{fro}^2}{\|A\|_2^2}$ is a smoothed version of the algebraic rank $rank(A)$. It is always true that $1 \leq \rho(A) \leq Rank(A) \leq \min(m, n)$. If $\rho(A) \leq 1 + \varepsilon$ for a sufficiently small ε the matrix is "close" to being of rank 1. Give an expression to the numerical rank of A in terms of its singular values σ_i. Express the numerical rank of $(AA^T)^k A$ in term of σ_i.

4. Assume that the matrix A is such that $\sigma_2/\sigma_1 \leq \eta$ for some $\eta < 1$. Use your expressions from above to find k such that $\rho((AA^T)^k A)) \leq 1 + \varepsilon$. How does this relate to the the Power Method for computing the largest singular value and vectors of A?