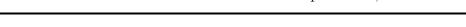
Prof. D. Kressner

H. Lam



## 1 ▶ Johnson-Lindenstrauss property for sub-Gaussian matrices

Let A be a  $\mathbb{R}^{n \times d}$  random matrix whose entries are independent, symmetric, mean zero, sub-Gaussian(1) and let  $Q = \frac{1}{\sqrt{n}}A$ .

- a) Prove that Q is isotropic, i.e., given a fixed vector  $x \in \mathbb{R}^d$ , we have  $\mathbb{E}[\|Qx\|_2^2] = \|x\|_2^2$ .
- b) Prove that if v is symmetric mean-zero sub-Gaussian( $\sigma$ ) random variable, and  $w \sim N(0, \sigma^2)$ , then  $\mathbb{E}[e^{tv^2}] \leq \mathbb{E}[e^{tw^2}]$  for t > 0.

Hint: To prove this bound, consider the proof technique used in Exercise 3.

c) Using the Chernoff bound, give a bound on n such that  $Q = \frac{1}{\sqrt{m}}A$  satisfies the Johnson-Lindenstrauss property. What about Oblivious Subspace Embedding property for a k-dimensional subspace?

## 2 ► Randomized least-squares problem

In this exercise, we study a classic iterative method for solving the for least-squares problem. The conjugate gradient (CG) method is usually introduced as a technique for solving positive definite linear systems Gx = h where  $G \in \mathbb{R}^{n \times n}$  is symmetric positive definite matrix. The advantage of CG over matrix factorization methods is that CG only requires access to the action  $x \to Gx$ .

a) Explain why solving

$$\min\{\|Ax - b\|_2 : x \in \mathbb{R}^m\}, \quad A \in \mathbb{R}^{d \times m}, b \in \mathbb{R}^d, \tag{1}$$

is equivalent to solving the linear system  $A^TAx = A^Tb$ .

- b) Study the conjugate gradient (CG) method in Algorithm 8.1 and Theorem 8.3 in Tropp'2020.
- c) Solve the linear system using the CG and preconditioned CG methods with incomplete Cholesky factorization as the preconditioner. The matrix  $A \in \mathbb{R}^{2000 \times 100}$  and  $b \in \mathbb{R}^{2000}$  can be found on Moodle. Plot the relative error vs. iterations.

Hint: you can use matlab command pcg and ichol.

d) We can also use a sketch to build a preconditioner, suppose S is an  $(m, \epsilon, \delta)$ -Oblivious Subspace Embedding, we can take  $R^{-1}$  as a preconditioner, where QR = SA is a QR factorization. Prove that

$$(1 - \epsilon)A^T A \le R^T R \le (1 + \epsilon)A^T A$$

and the condition number of  $(AR^{-1})^TAR^{-1}$  is bounded by  $(1 + \epsilon)/(1 - \epsilon)$  with probability  $1 - \delta$ .

e) Use Gaussian random matrix and subsampled cosine transform matrix as S to build a preconditioner. Plot the relative error vs. iterations.

Hint: You can use matlab command dct to apply the cosine transform.

## 3 ► Sparse sign matrices

Let  $U \in \mathbb{R}^{d \times k}$  be an orthonormal basis. Let S be a  $n \times d$  i.i.d. sparse sign matrix defined as in the Lecture 5 slide 16.

- a) Prove that S is isotrpic.
- b) Denote  $s_j^T$  the jth row of S and  $X_j := \frac{1}{pn} U^T s_j s_j^T U$ . Show that

$$\mathbb{E}||X_j||_2 = \frac{k}{n},$$

and provide an upper bound for  $\|X_j\|_2$ . Is the bound you provided tight? Construct an example to illustrate that.