Assignment 2

Randomization in Numerical Linear Algebra (PCMI)

1. Using the notation of slide 17, prove that

$$\mathbf{E}\left[(\mathbf{CR})_{ij}\right] = (\mathbf{AB})_{ij}\,.$$

2. Using the notation of slide 17, prove that

$$\mathbf{Var}\left[(\mathbf{CR})_{ij}\right] = \frac{1}{c}\sum_{k=1}^{n}\frac{\mathbf{A}_{ik}^2\mathbf{B}_{kj}^2}{p_k} - \frac{1}{c}(\mathbf{AB})_{ij}^2\,.$$

3. Using the notation of slide 19 (and the sampling probabilities of slide 13[1]), prove that

$$\mathbf{E}\left[\,\|\mathbf{AB}-\mathbf{CR}\|_F^2\right] \le \frac{1}{c}\|\mathbf{A}\|_F^2\,\|\mathbf{B}\|_F^2\,.$$

Then, apply Markov's inequality to get a bound that holds with probability at least 0.9.

4. In this problem, we will complete the proof of the spectral norm bound of the error matrix $\mathbf{AA}^T - \mathbf{CC}^T$ (using the assumptions of slide 23). We will use the notation introduced at slides 23-26.

1. First, using the notation of slide 25, prove

$$\mathbf{E}\left[\mathbf{yy}^T\right] = \mathbf{AA}^T \quad \text{and} \quad \mathbf{CC}^T = \frac{1}{c}\sum_{t=1}^{c}\mathbf{y}^t(\mathbf{y}^t)^T.$$

2. Second, using the bounds on M and the $\left\|\mathbf{E}\left[\mathbf{yy}^T\right]\right\|_2$, apply the matrix concentration inequality of slide 26 to prove that

$$\left\|\mathbf{AA}^T - \mathbf{CC}^T\right\|_2 \le \varepsilon,$$

with probability at least $1-\delta$ by setting

$$c = \Omega\left(\frac{\|\mathbf{A}\|_F^2}{\epsilon^2}\ln\left(\frac{\|\mathbf{A}\|_F^2}{\epsilon^2\sqrt{\delta}}\right)\right).$$

[1] Why were those particular probabilities chosen in the algorithm?