Homework 7. Due Wednesday, Oct. 16.

Dataset: An incomplete spreadsheet of movie ratings. Data file: `MovieRankingData2024.csv`. If you haven't done it, please feel free to manually add your own row there. Format: CSV (can be opened and edited e.g. using Numbers (Mac OS), Excel (Windows)).

Programming: Pick any language you wish. High-level language, e.g. Matlab or Python, is preferable. All requested algorithms should be programmed from scratch. Please use standard functions for SVD.

1. **(10 pts)** Do **matrix completion** in two ways:

   (a) Use the low-rank factorization model $A \approx XY^\top$ and the objective function of the form

   $$F(X,Y) = \frac{1}{2}\|P_\Omega(A - XY^\top)\|_F^2 + \frac{\lambda}{2}\left(\|X\|_F^2 + \|Y\|_F^2\right).$$

   Try values of λ 0.1, 1, and 10, and $\mathsf{rank}(X) = \mathsf{rank}(Y) = k$, $k = 1, 2, \ldots, 7$. Find X and Y using alternating iteration

   $$X^{m+1} = \arg\min_X F(X, Y^m), \tag{1}$$
   $$Y^{m+1} = \arg\min_Y F(X^{m+1}, Y). \tag{2}$$

   Each of these steps can be further decomposed into a collection of small linear least squares problems. For example, at each substep of (1), we solve the linear least squares problem to compute the row i of X:

   $$\mathbf{x}_i^\top = \arg\min_{\mathbf{x}} \frac{1}{2}\left\|\mathbf{x}^\top Y_{\Omega_i}^\top - a_{\Omega_i}\right\|_2^2 + \frac{\lambda}{2}\|\mathbf{x}\|^2, \tag{3}$$

   where $\Omega_i := \{j \mid (i,j) \in \Omega\}$, $Y_{\Omega_i}^\top$ is the set of columns of $Y^\top$ with indices in Ω_i, and a_{Ω_i} is the set of known entries of A in its row i. A similar problem can be set up for each column of Y. Work out solutions to these problems in a manner similar to the one in Sections 5.3.1 and 5.3.2 of `LinearAlgebra.pdf` except that there will be no constraint requiring the entries to be positive. Implement the resulting algorithm. Comment on how the value of λ and the choice of rank affects the result. Which values of the rank and λ seem the most reasonable to you? You can judge by your own row.

   (b) Use the approach of penalizing the nuclear norm in Section 6.3 of `LinearAlgebra.pdf` and the iteration

   $$M^{j+1} = S_\lambda(M^j + P_\Omega(A - M^j))$$

   Experiment with different values of λ.

Compare these two approaches for matrix completion. Which one gives more sensible results? Which one is easier to use? Which one do you find more efficient?

2. **(5 pts)** Adding the nuclear norm into the objective function for the matrix completion problem promotes low rank. The reason is that level sets of $\|A\|_* = a$ tend to have corners, and the minimizers of the objective function including the nuclear norm penalty term tend to lie at these corners. Check experimentally that nuclear norm level set for a 2×2 matrix

$$A = \begin{bmatrix} 0.5 & x \\ y & z \end{bmatrix}$$

have corners at points where $\det A = 0$. Proceed as follows. Let $a = 1$. Set a grid for the values of z:

$$z \in \{-0.4, -0.3, -0.2, -0.1, 0, 0.1, 0.2, 0.3, 0.4\}.$$

For each grid value of z plot two curves on the same figure. Curve 1: the zero level set of the determinant of A. Curve 2: the level set $\|A\|_* = a$. Note at which points these curves intersect.

3. **(3 pts)** Let A be a symmetric positive definite matrix and let $\{p_0, \dots, p_{n-1}\}$ be a set of vectors such that $p_j^\top A p_k = 0$ for all $j \neq k$, $0 \leq j, k \leq n-1$. We say that this set of vectors *conjugate* with respect to the matrix A. Prove that this set of vectors must be linearly independent.