CSE250B Homework 7

Qiao Zhang

March 6, 2016

1 Singular value versus eigenvalue

1.1 a

$$Mv_i = \begin{bmatrix} | & ... & | \\ u_1 & ... & u_p \\ | & ... & | \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & ... & 0 \\ 0 & 0 & \sigma_p \end{bmatrix} \begin{bmatrix} - & v_1 & - \\ - & ... & - \\ - & v_p & - \end{bmatrix} v_i = \sigma_i u_i$$

1.2 b

$$M^T u_i = \begin{bmatrix} | & ... & | \\ v_1 & ... & v_p \\ | & ... & | \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & ... & 0 \\ 0 & 0 & \sigma_p \end{bmatrix} \begin{bmatrix} - & u_1 & - \\ - & ... & - \\ - & u_p & - \end{bmatrix} u_i = \sigma_i v_i$$

1.3 c

$$M^T M v_i = M^T \sigma_i u_i = \sigma_i^2 v_i$$
$$M M^T u_i = M \sigma_i v_i = \sigma_i^2 u_i$$

1.4 d

For MM^T, σ_i^2 are eigenvalues and u_i are eigenvectors.

1.5 e

For $M^T M$, σ_i^2 are eigenvalues and v_i are eigenvectors.
MM^T and $M^T M$ share the same eigenvalues while their eigenvectors are u_i and v_i respectively.

$$M^T M = V(S^T S)V^T MM^T = U(SS^T)U^T$$

1.6 f

If $\sigma_1^2 \geq \sigma_2^2 \geq ... \geq \sigma_p^2$, then $\forall i > k, \sigma_k = 0$.

2 Rank-1 matrices

2.1 a

$$M = USV^T$$

$$U = \begin{bmatrix} -0.39 & -0.92 \\ -0.92 & 0.39 \end{bmatrix} S = \begin{bmatrix} 9.51 & 0 \\ 0 & 0.77 \end{bmatrix} V^T = \begin{bmatrix} -0.43 & -0.57 & -0.70 \\ -0.81 & 0.11 & -0.58 \end{bmatrix}$$

$$\hat{M} = \begin{bmatrix} 1.57 & 2.08 & 2.59 \\ 3.76 & 4.97 & 6.17 \end{bmatrix}$$

2.2 b

The decomposition is not unique.
For any orthonormal matrix Q, such that $Q^TQ = QQ^T = I$, we have

$$M = USV^T = UQQ^TSQQ^TV^T = (UQ)(Q^TSQ)(VQ)^T$$

thus,

$$ab^T = (uQ)(vQ)^T = uv^T$$

2.3 c

$$\hat{M} = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + ... + \sigma_k u_k v_k^T$$

3 Gram matrix

$$G = X^T X = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 2 \\ 1 & 1 & 2 & 2 \\ 1 & 2 & 2 & 3 \end{bmatrix}$$

There are other sets of four points with exactly the same Gram matrix. For example, if Y = -X, then gram(X) = gram(Y) = $X^T X$.

4 Classical multidimensional scaling

The result is shown below.
I think it makes some sense but still needs some improvement. In order to obtain a better embedding, I would add more reasonable features and slightly tune the parameters of the MDS model such as max_iteration and eps.
I am not that familiar with metropolis in America, but I would guess Chicago as Mystery because it has similar performance with New York, Atlanta and Denver in various rankings such as population and public security. It also appears most frequent on the web page once I Google New York, Atlanta and Denver.

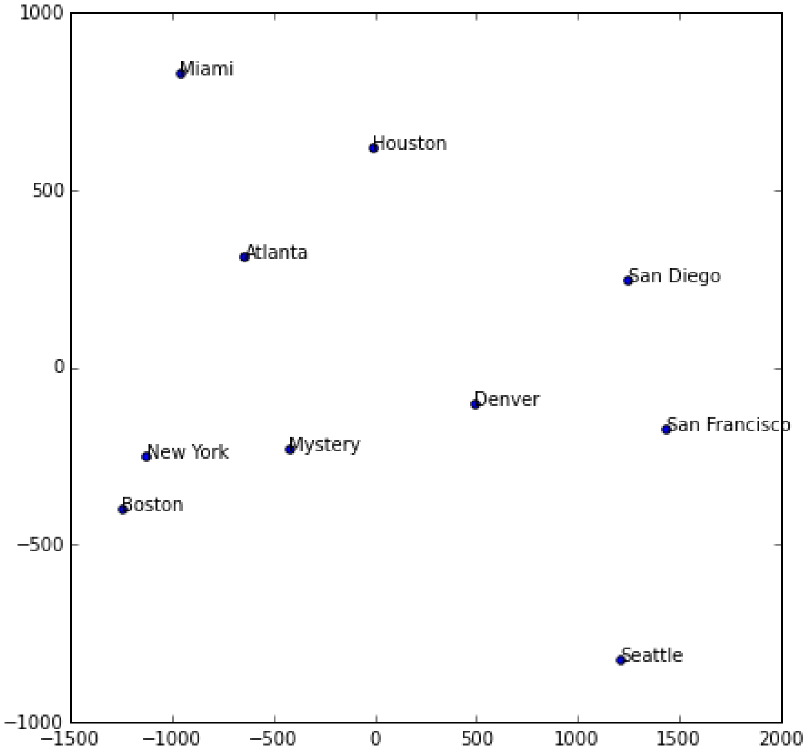

Figure 1: MDS plot