MATH 156S HOMEWORK 9

Due Jun. 4

Exercise 1. The topics learned by NMF in Figures 52 and 53 uses tf-idf vectorizer of the documents. Use bag-of-words vectorizer instead and reproduce these figures. Can you see any difference? Use the code in Use the code in Jupyter Notebook that generates these plots.

Exercise 2 (Projection)**.** Let $\mathbf{x} \in \mathbb{R}^d$ and $C \subseteq \mathbb{R}^d$. The projection $\text{Proj}_C(\mathbf{x})$ of $\mathbf{x}$ onto C is defined by

$$\text{Proj}_C(\mathbf{x}) := \underset{\mathbf{u}\in C\subseteq\mathbb{R}^d}{\arg\min}\ \|\mathbf{x}-\mathbf{u}\|_F^2. \tag{1}$$

Namely, $\text{Proj}_C(\mathbf{x})$ is the point in C that is closest to $\mathbf{x}$. It is defined as a constrianed quadratic minimization as above.

(i) Write $\|\mathbf{x}-\mathbf{u}\|_F^2 = \text{tr}((\mathbf{x}-\mathbf{u})^T(\mathbf{x}-\mathbf{u})) = \text{tr}(\mathbf{u}^T\mathbf{u}) - 2\text{tr}(\mathbf{x}^T\mathbf{u}) + \text{tr}(\mathbf{x}^T\mathbf{x})$. Show that

$$\frac{\partial}{\partial \mathbf{u}}\|\mathbf{x}-\mathbf{u}\|_F^2 = 2(\mathbf{u}-\mathbf{x}). \tag{2}$$

(ii) Assume C is the one-dimensional span $\langle \mathbf{v}\rangle := \{a\mathbf{v}\,|\, a\in\mathbb{R}\}$. Then

$$\text{Proj}_{\mathbf{v}}(\mathbf{x}) := \text{Proj}_{\langle\mathbf{v}\rangle}(\mathbf{x}) = \underset{a\in\mathbb{R}}{\arg\min}\ \|\mathbf{x}-a\mathbf{v}\|_F^2. \tag{3}$$

Use **(i)** and Lagrange multiplier or solve the above unconstrained optimization directly and show

$$\text{Proj}_{\mathbf{v}}(\mathbf{x}) = (\mathbf{x}^T\mathbf{v})\mathbf{v}. \tag{4}$$

(iii) Let $\mathbf{W} = [\mathbf{u}_1, \dots, \mathbf{u}_r] \in \mathbb{R}^{d\times r}$ be such that its columns are orthonormal, i.e., $\mathbf{W}^T\mathbf{W} = \mathbf{I} \in \mathbb{R}^{r\times r}$. Show that

$$\text{Proj}_{\langle\mathbf{W}\rangle}(\mathbf{x}) := \text{Proj}_{\langle\mathbf{u}_1,\dots,\mathbf{u}_r\rangle}(\mathbf{x}) = \sum_{i=1}^{r}(\mathbf{u}_i^T\mathbf{x})\mathbf{u}_i = \mathbf{W}(\mathbf{W}^T\mathbf{x}). \tag{5}$$

Exercise 3 (Cyclic property of trace)**.** Show that $\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{BCA})$.

Exercise 4 (Spectral Theorem for Real Symmetric Matrices)**.** Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a real symmetric matrix, i.e., $\mathbf{A}^T = \mathbf{A}$. Then the spectral theorem for real symmetric matrices states that the following hold:

(1). $\mathbf{A}$ has n real eigenvalues (counting multiplicities);

(2). There exists an orthonormal matrix $\mathbf{U} \in \mathbb{R}^{n\times n}$, $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ and a diagonal matrix $\mathbf{D} \in \mathbb{R}^{n\times n}$ such that

$$\mathbf{A} = \mathbf{U}^T\mathbf{D}\mathbf{U}; \tag{6}$$

(3). In **(ii)**, the orthonomal matrix $\mathbf{U}$ consists of eigenvectors of $\mathbf{A}$ and the diagonal entries of $\mathbf{D}$ are the eigenvalues of $\mathbf{A}$;

In this exercise, we will prove this result by an induction on n.

(i) Show that all of the eigenvalues of $\mathbf{A}$ are real.

(ii) Let $\mathbf{v}_1$ be an eigenvector of $\mathbf{A}$ and let λ_1 be its associated eigenvalue. By **(i)**, λ_1 is real. Extend $\mathbf{v}_1$ to an orthonormal basis $\mathcal{B} := \{\mathbf{v}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ for $\mathbb{R}^n$ and let $\mathbf{Q} := [\mathbf{v}_1, \mathbf{u}_2, \dots, \mathbf{u}_n] \in \mathbb{R}^{n\times n}$. Show that

$$\mathbf{Q}^T\mathbf{A}\mathbf{Q} = \left[\begin{array}{c|ccc} \lambda_1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & \mathbf{A}' & \\ 0 & & & \end{array}\right], \tag{7}$$

where $\mathbf{A}' \in \mathbb{R}^{(n-1)\times(n-1)}$ is a real symmetric matrix.

(iii) Suppose there exists an orthonormal matrix $\mathbf{V} \in \mathbb{R}^{(n-1)\times(n-1)}$ and a diagonal matrix $\mathbf{D}' \in \mathbb{R}^{(n-1)\times(n-1)}$ such that $\mathbf{A}' = \mathbf{V}^T\mathbf{D}'\mathbf{V}$. Define

$$\overline{\mathbf{V}} = \left[\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & \mathbf{V} & \\ 0 & & & \end{array}\right] \in \mathbb{R}^{n\times n}. \tag{8}$$

Then show that

$$\mathbf{A} = \mathbf{Q}\overline{\mathbf{V}}^T \left[\begin{array}{c|ccc} \lambda_1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & \mathbf{D}' & \\ 0 & & & \end{array}\right] \overline{\mathbf{V}}\mathbf{Q}^T. \tag{9}$$

Also show that $\mathbf{U} := \overline{\mathbf{V}}\mathbf{Q}^T \in \mathbb{R}^{n\times n}$ is orthonomal.

(iii) By using **(i)**-**(ii)**, prove that if $\mathbf{A} \in \mathbb{R}^{n\times n}$ is symmetric, show that there are orthogonal and diagonal matrices $\mathbf{U}, \mathbf{D} \in \mathbb{R}^{n\times n}$ such that $\mathbf{A} = \mathbf{U}^T\mathbf{D}\mathbf{U}$ and $\mathbf{D}$ consists of the eigenvalues of $\mathbf{A}$ as its diagonal entries. Also deduce that the columns of $\mathbf{U}$ are the eigenvectors of $\mathbf{A}$.