Data Analysis and Matrix Computations Assignment 2

Exercise 1.

Derive the following formulas in Lecture 6

$$\frac{\partial f(\mathbf{W}, \mathbf{H})}{\partial \zeta_{\mu\nu}} = -\mathbf{H}_{\mu\nu}(\mathbf{W}^\top \mathbf{X})_{\mu\nu} + \mathbf{H}_{\mu\nu}(\mathbf{W}^\top \mathbf{W}\mathbf{H})_{\mu\nu},$$

and

$$\frac{\partial g(\mathbf{W}, \mathbf{H})}{\partial \zeta_{\mu\nu}} = \sum_i \left(-\frac{\mathbf{X}_{i\nu}}{(\mathbf{W}\mathbf{H})_{i\nu}} + 1 \right) \mathbf{W}_{i\mu}\mathbf{H}_{\mu\nu}.$$

Exercise 2.

Prove the following: For any two matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times p}_{++}$, the entropy divergence, defined as

$$D(\mathbf{A}||\mathbf{B}) = \sum_{i=1}^{n} \sum_{j=1}^{p} \left(\mathbf{A}_{ij} \log \frac{\mathbf{A}_{ij}}{\mathbf{B}_{ij}} - \mathbf{A}_{ij} + \mathbf{B}_{ij} \right),$$

is nonnegative, and is equal to zero if and only if $\mathbf{A} = \mathbf{B}$.

Exercise 3.

Let

$$\pi_{\mathcal{C}}(\mathbf{x}) := \operatorname*{argmin}_{\mathbf{y}\in\mathcal{C}} \|\mathbf{x} - \mathbf{y}\|_2$$

denote the projection of $\mathbf{x} \in \mathbb{R}^n$ onto the set $\mathcal{C} \subseteq \mathbb{R}^n$. Prove the following: If $\mathcal{C}$ is closed convex, then for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\|\pi_{\mathcal{C}}(\mathbf{x}) - \pi_{\mathcal{C}}(\mathbf{y})\|_2 \leq \|\mathbf{x} - \mathbf{y}\|_2.$$

Exercise 4.

If

$$\mathbb{S}_n = \{\mathbf{A} \in \mathbb{R}^{n\times n} : \mathbf{A} = \mathbf{A}^\top\}$$

is the set of symmetric matrices and, for $\mathbf{A} \in \mathbb{S}_n$, $f : \mathbb{S}_n \mapsto \mathbb{R}$ given by $f(\mathbf{A}) = \lambda_{\max}(\mathbf{A})$ (maximum eigenvalue of $\mathbf{A}$), show that f is convex and find a rank-one matrix in $\partial f(\mathbf{A})$. Hint: (1) for convexity, show $\mathbb{S}_n$ is convex and f satisfies Jensen's inequality; (2) for subgradient, the involved inner product is $\langle \mathbf{A}, \mathbf{B} \rangle = \mathrm{tr}(\mathbf{A}^\top \mathbf{B})$.

Exercise 5.

Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$ be given. Compute $\partial\|\mathbf{A}\mathbf{x} + \mathbf{b}\|_1$, $\partial\|\mathbf{A}\mathbf{x} + \mathbf{b}\|_2$ and $\partial\|\mathbf{A}\mathbf{x} + \mathbf{b}\|_\infty$.

Exercise 6.

Given a transition matrix $\mathbf{P} \in \mathbb{R}^{N\times N}$. For $0 < \varepsilon < 1/N$, define

$$\mathbf{P}_\varepsilon = (1 - N\varepsilon)\mathbf{P} + \varepsilon \mathbf{1}\mathbf{1}^\top.$$

Assume that $\mathbf{P}$ has eigenvalues $\{1, \lambda_2, \cdots, \lambda_N\}$. Prove that $\mathbf{P}_\varepsilon$ has eigenvalues $\{1, \alpha\lambda_2, \cdots, \alpha\lambda_N\}$, where $\alpha = (1 - N\varepsilon) < 1$.