Manifold Learning and Sparse Representation: Assignment 11

ZHANG Yuan, 1601111332

Exercise 194.

Proof. Prove by verifying the following:

- By definition,

$$\Omega^G_{overlap}(w) = \inf_{v \in V_G, \sum_{g\in G} v_g = w} \sum_{g \in G} \|v_g\| \geq 0,$$

  and the equation holds iff. there exists a v such that $\|v_g\| = 0$ for each g, which deduces to $v_g = 0$, and hence iff. $w = \sum_{g\in G} v_g = 0$;

- By definition,

$$\Omega^G_{overlap}(\alpha w) = \inf_{v \in V_G, \sum_{g\in G} v_g = \alpha w} \sum_{g \in G} \|v_g\| = \alpha \inf_{v \in V_G, \sum_{g\in G} \frac{1}{\alpha} v_g = w} \sum_{g \in G} \|\frac{1}{\alpha} v_g\| = \alpha \Omega^G_{overlap}(w);$$

- We have

$$\begin{aligned}
\Omega^G_{overlap}(w_1) + \Omega^G_{overlap}(w_2) &= \inf_{v_1 \in V_G, \sum_{g\in G} v_{1,g} = w_1} \sum_{g \in G} \|v_{1,g}\| + \inf_{v_2 \in V_G, \sum_{g\in G} v_{2,g} = w_2} \sum_{g \in G} \|v_{2,g}\| \\
&= \inf_{v_1 \in V_G, v_2 \in V_G, \sum_{g\in G} v_{g,1} = w_1, \sum_{g\in G} v_{g,2} = w_2} \sum_{g \in G} (\|v_{1,g}\| + \|v_{2,g}\|) \\
&\geq \inf_{v_1 \in V_G, v_2 \in V_G, \sum_{g\in G} v_{g,1} = w_1, \sum_{g\in G} v_{g,2} = w_2} \sum_{g \in G} \|v_{1,g} + v_{2,g}\| \\
&= \inf_{v_1 + v_2 \in V_G, \sum_{g\in G} v_{g,1} = w_1, \sum_{g\in G} v_{g,2} = w_2} \sum_{g \in G} \|v_{1,g} + v_{2,g}\| \\
&\geq \inf_{v \in V_G, \sum_{g\in G} v = w_1 + w_2} \sum_{g \in G} \|v\| = \Omega^G_{overlap}(w_1 + w_2),
\end{aligned}$$

  where the last inequality holds thanks to the fact that $v_1 + v_2 \in G$ and $\sum_{g\in G} v_{g,1} + v_{g,2} = w_1 + w_2$.

□

Exercise 195.

Proof. According to Cauchy-Schwarz's inequality, assuming $\|\alpha_g\| \leq 1$,

$$\begin{aligned}\alpha^T w = \alpha^T \sum_g v_g &= \sum_g \alpha_g^T v_g \\ &\leq \sum_g \|\alpha_g\| \cdot \|v_g\| \\ &\leq \sum_g \|v_g\|,\end{aligned}$$

Thus,

$$\Omega^G_{overlap}(w) = \inf_{v \in V_G, \sum_{g\in G} v_g = w} \sum_{g\in G} \|v_g\| \geq \sup_\alpha \{\alpha^T w | \|\alpha_g\| \leq 1\},$$

where the equation can be obtained when v_g and α_g are linear dependent (a special case of which is nonoverlapping $\{v_g\}$ with corresponding α). □

Exercise 196.

Suppose we have F levels of frequencies, i.e., F anti-diagonals. We can use the following penalty term,

$$\sum_{f=1}^{F} \|\{D_{i,j} | i + j \leq f\}\|.$$

That is, for each f, we sum up all the entries that is upper left to the f-th anti-diagonal.

Exercise 197.

Because when there is overlapping between groups, the covariates that belong to multiple groups are penalized multiple times in the group lasso penalty and hence only the sparsity of non-overlapping covariates (i.e., w_1, w_3).

On the contrary, the overlapping penalty is able to encourage sparsity of each group of covariates by decomposing the overlapping part (i.e., w_2) into their corresponding groups.

Exercise 200.

Proof. According to Proposition 199,

$$Z^* = V V_O^T = \begin{bmatrix} V_O \\ V_H \end{bmatrix} \cdot V_O^T = \begin{bmatrix} V_O V_O^T \\ V_H V_O^T \end{bmatrix}.$$

Then plug into Eq. (10.19),

$$\begin{aligned} D &= \begin{bmatrix} D & D_H \end{bmatrix} \cdot \begin{bmatrix} V_O V_O^T \\ V_H V_O^T \end{bmatrix} \\ &= D V_O V_O^T + D_H V_H V_O^T \\ &= D V_O V_O^T + D_H V_H \Sigma^{\dagger} U^T (U \Sigma V_O^T) \\ &= D V_O V_O^T + D_H V_H \Sigma^{\dagger} U^T D \\ &= DZ *_{O|H} + L^*_{H|O} D, \end{aligned}$$

where $Z^*_{O|H} \triangleq V_O V_O^T$ and $L^*_{H|O} = D_H V_H \Sigma^{\dagger} U^T$ (note that typically Σ is not full rank, otherwise $L^* = I$). □

Exercise 202.

Proof. Prove by verifying the following:

- By definition, $\|D\text{Diag}(w)\|_* \geq 0$. The equation holds iff. $D\text{Diag}(w) = 0$, which is equivalent to $D_{:i} w_i = 0$ for each i, that is, $w = 0$ because none of $D_{:i} = 0$;
- $\|D\text{Diag}(\alpha w)\|_* = \|\alpha D\text{Diag}(w)\|_* = \alpha \|D\text{Diag}(w)\|_*$;
- Thanks to the definition of norms,

$$\begin{aligned} \|D\text{Diag}(w_1 + w_2)\|_* &= \|D\text{Diag}(w_1) + D\text{Diag}(w_2)\|_* \\ &\leq \|D\text{Diag}(w_1)\|_* + \|D\text{Diag}(w_2)\|_* \end{aligned}$$

□