Appendices

Supplementary Materials for "**Identifying diagnosis-specific genotype-phenotype associations via joint multi-task sparse canonical correlation analysis and classification**" by Lei Du, Fang Liu, Kefei Liu, Xiaohui Yao, Shannon L. Risacher, Junwei Han, Lei Guo, Andrew J. Saykin and Li Shen, for the Alzheimer's Disease Neuroimaging Initiative.

A Convergence analysis

We have the following theorem for the MT-SCCALR Algorithm.

Theorem 1. *The Algorithm 1 decreases the objective in each iteration.*

Proof. Since we need to optimize $\mathbf{V}$ and $\mathbf{U}$ alternatively, both the convergence of Eq. (13) and Eq. (19) should be provided.

Firstly, we use $\hat{\mathbf{V}}$ to denote the updated $\mathbf{V}$, $\hat{\mathbf{v}}_c$ to denote the updated $\mathbf{v}_c$ and so on. From Eq. (18), we have

$$\begin{aligned}
&\mathcal{L}_{LR}(\hat{\mathbf{V}}) + 2\sum_{c=1}^{C} \hat{\mathbf{v}}_c^\top \mathbf{Y}_c^\top \mathbf{X}_c \mathbf{u}_c + \lambda_{v1} Tr(\hat{\mathbf{V}}^\top \hat{\tilde{D}}_1 \hat{\mathbf{V}}) + \lambda_{v2} Tr(\hat{\mathbf{V}}^\top \hat{\bar{\mathbf{D}}}_1 \hat{\mathbf{V}}) \\
&+ \lambda_{v3} Tr(\hat{\mathbf{V}}^\top \hat{\mathbf{D}}_1 \hat{\mathbf{V}}) + (\gamma_v + 1) \sum_{c=1}^{C} \|\mathbf{Y}_c \hat{\mathbf{v}}_c\|_2^2 \\
&\leq \mathcal{L}_{LR}(\mathbf{V}) + 2\sum_{c=1}^{C} \mathbf{v}_c^\top \mathbf{Y}_c^\top \mathbf{X}_c \mathbf{u}_c + \lambda_{v1} Tr(\mathbf{V}^\top \tilde{\mathbf{D}}_1 \mathbf{V}) + \lambda_{v2} Tr(\mathbf{V}^\top \bar{\mathbf{D}}_1 \mathbf{V}) \\
&+ \lambda_{v3} Tr(\mathbf{V}^\top \mathbf{D}_1 \mathbf{V}) + (\gamma_v + 1) \sum_{c=1}^{C} \|\mathbf{Y}_c \mathbf{v}_c\|_2^2 \Longrightarrow \mathcal{L}_{LR}(\hat{\mathbf{V}}) \\
&+ \sum_{c=1}^{C} \|\mathbf{Y}_c \hat{\mathbf{v}}_c - \mathbf{X}_c \mathbf{u}_c\|_2^2 + \lambda_{v1} \sum_{j=1}^{q} \frac{\|\hat{\mathbf{v}}^j\|_2^2}{2\|\mathbf{v}^j\|_2} + \lambda_{v2} \sum_{j=1}^{q}\sum_{c=1}^{C} \frac{\|\hat{v}_{jc}\|_2^2}{2\|v_{jc}\|_2} \\
&+ \lambda_{v3} \sum_{c=1}^{C} \sum_{j=1,(j,k)\in E}^{q} \frac{\|\hat{v}_{jc}\|_2^2}{2\sqrt{v_{jc}^2 + v_{kc}^2}} + \gamma_v \sum_{c=1}^{C} \|\mathbf{Y}_c \hat{\mathbf{v}}_c\|_2^2 \leq \mathcal{L}_{LR}(\mathbf{V}) \\
&+ \sum_{c=1}^{C} \|\mathbf{Y}_c \mathbf{v}_c - \mathbf{X}_c \mathbf{u}_c\|_2^2 + \lambda_{v1} \sum_{j=1}^{q} \frac{\|\mathbf{v}^j\|_2^2}{2\|\mathbf{v}^j\|_2} + \lambda_{v2} \sum_{j=1}^{q}\sum_{c=1}^{C} \frac{\|v_{jc}\|_2^2}{2\|v_{jc}\|_2} \\
&+ \lambda_{v3} \sum_{c=1}^{C} \sum_{j=1,(j,k)\in E}^{q} \frac{\|v_{jc}\|_2^2}{2\sqrt{v_{jc}^2 + v_{kc}^2}} + \gamma_v \sum_{c=1}^{C} \|\mathbf{Y}_c \mathbf{v}_c\|_2^2 .
\end{aligned} \tag{1}$$

According to the Lemma 1 in (Wang *et al.*, 2012b; Nie *et al.*, 2010), we have $\|\hat{v}_{jc}\|_2 - \frac{\|\hat{v}_{jc}\|_2^2}{2\|v_{jc}\|_2} \leq \|v_{jc}\|_2 - \frac{\|v_{jc}\|_2^2}{2\|v_{jc}\|_2}$ and $\|\hat{\mathbf{v}}^j\|_2 - \frac{\|\hat{\mathbf{v}}^j\|_2^2}{2\|\mathbf{v}^j\|_2} \leq \|\mathbf{v}^j\|_2 - \frac{\|\mathbf{v}^j\|_2^2}{2\|\mathbf{v}^j\|_2}$. Similarly, we can derive $\sqrt{\hat{v}_{jc}^2 + \hat{v}_{jk}^2} - \frac{\hat{v}_{jc}^2 + \hat{v}_{jk}^2}{2\sqrt{v_{jc}^2 + v_{jk}^2}} \leq \sqrt{v_{jc}^2 + v_{jk}^2} - \frac{v_{jc}^2 + v_{jk}^2}{2\sqrt{v_{jc}^2 + v_{jk}^2}}$.

Then plugging them into Eq. (1) for multiple times regarding each feature respectively, we arrive at

$$\begin{aligned}
&\mathcal{L}_{LR}(\hat{\mathbf{V}}) + \sum_{c=1}^{C} \|\mathbf{Y}_c \hat{\mathbf{v}}_c - \mathbf{X}_c \mathbf{u}_c\|_2^2 + \lambda_{v1} \sum_{j=1}^{q} \left\|\hat{\mathbf{v}}^j\right\|_2 + \lambda_{v2} \sum_{j=1}^{q}\sum_{c=1}^{C} \|\hat{v}_{jc}\|_2 \\
&+ \lambda_{v3} \sum_{c=1}^{C} \sum_{j=1,(j,k)\in E}^{q} \sqrt{\hat{v}_{jc}^2 + \hat{v}_{kc}^2} + \gamma_v \sum_{c=1}^{C} \|\mathbf{Y}_c \hat{\mathbf{v}}_c\|_2^2 \leq \mathcal{L}_{LR}(\mathbf{V}) \\
&+ \sum_{c=1}^{C} \|\mathbf{Y}_c \mathbf{v}_c - \mathbf{X}_c \mathbf{u}_c\|_2^2 + \lambda_{v1} \sum_{j=1}^{q} \left\|\mathbf{v}^j\right\|_2 + \lambda_{v2} \sum_{j=1}^{q}\sum_{c=1}^{C} \|v_{jc}\|_2 \\
&+ \lambda_{v3} \sum_{c=1}^{C} \sum_{j=1,(j,k)\in E}^{q} \sqrt{v_{jc}^2 + v_{kc}^2} + \gamma_v \sum_{c=1}^{C} \|\mathbf{Y}_c \mathbf{v}_c\|_2^2 .
\end{aligned}$$

Writing this inequations into matrix form yields

$$\begin{aligned}
&\mathcal{L}_{LR}(\hat{\mathbf{V}}) + \sum_{c=1}^{C} \|\mathbf{Y}_c \hat{\mathbf{v}}_c - \mathbf{X}_c \mathbf{u}_c\|_2^2 + \lambda v1 \left\|\hat{\mathbf{V}}\right\|_{2,1} + \lambda_{v2} \left\|\hat{\mathbf{V}}\right\|_{1,1} \\
&+ \lambda_{v3} \sum_{c=1}^{C} \|\hat{\mathbf{v}}_c\|_{GGL} + \gamma_v \sum_{c=1}^{C} \|\mathbf{Y}_c \hat{\mathbf{v}}_c\|_2^2 \leq \mathcal{L}_{LR}(\mathbf{V}) \\
&+ \sum_{c=1}^{C} \|\mathbf{Y}_c \mathbf{v}_c - \mathbf{X}_c \mathbf{u}_c\|_2^2 + \lambda v1 \|\mathbf{V}\|_{2,1} + \lambda_{v2} \|\mathbf{V}\|_{1,1} \\
&+ \lambda_{v3} \sum_{c=1}^{C} \|\mathbf{v}_c\|_{GGL} + \gamma_v \sum_{c=1}^{C} \|\mathbf{Y}_c \mathbf{v}_c\|_2^2 .
\end{aligned}$$

Thus the objective decreases in each iteration.

Using the same method, we can prove that the objective also decreases in each iteration when optimizing $\mathbf{U}$. Then combining both conclusions together, we obtain the following inequations

$$\begin{aligned}
&\mathcal{L}_{LR}(\hat{\mathbf{U}}, \hat{\mathbf{V}}) + \mathcal{L}_{SCCA}(\hat{\mathbf{U}}, \hat{\mathbf{V}}) + \Omega(\hat{\mathbf{U}}) + \Omega(\hat{\mathbf{V}}) \\
&\leq \mathcal{L}_{LR}(\mathbf{U}, \mathbf{V}) + \mathcal{L}_{SCCA}(\mathbf{U}, \mathbf{V}) + \Omega(\mathbf{U}) + \Omega(\mathbf{V})
\end{aligned} \tag{2}$$

This finally completes the proof.

Acknowledgements

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. HoffmannLa Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.