Supplement: On Data Integration Problems with Manifolds

A Proofs

Proof of Lemma 1

Proof. The $\boldsymbol{y}_U$-score of objective

$$\sum_{i \in L} \phi\left(y_i, f_i\right) + \|\boldsymbol{y}_U - \boldsymbol{f}_U\|_2^2 + \lambda \boldsymbol{f}^T \boldsymbol{\Delta} \boldsymbol{f} + \gamma \|\boldsymbol{y}_U\|_2^2 \tag{1}$$

is $\boldsymbol{y}_U - \boldsymbol{f}_U + \gamma \boldsymbol{y}_U = \vec{0}$ and verifies $\boldsymbol{y}_U = (1 - p_\gamma) \boldsymbol{f}_U$ as optimal. Vector $\boldsymbol{y}_U$ can then be profiled out of (1) because

$$\gamma \|\boldsymbol{y}_U\|_2^2 + \|\boldsymbol{f}_U - \boldsymbol{y}_U\|_2^2 = p_\gamma \|\boldsymbol{f}_U\|_2^2 . \tag{2}$$

The $\boldsymbol{f}_U$-score is thus $\lambda \boldsymbol{\Delta}_{UL} \boldsymbol{f}_L + \lambda \boldsymbol{\Delta}_{UU} \boldsymbol{f}_U + p_\gamma \boldsymbol{f}_U = \vec{0}$ and profiling out $\boldsymbol{f}_U$ gives

$$\begin{aligned} \lambda \boldsymbol{f}_L^T \boldsymbol{\Delta}_{LL}^{\star} \boldsymbol{f}_L &= \lambda \boldsymbol{f}_L^T \boldsymbol{\Delta}_{LL} \boldsymbol{f}_L + 2\lambda \boldsymbol{f}_L^T \boldsymbol{\Delta}_{LU} + \boldsymbol{f}_U^T \left(\lambda \boldsymbol{\Delta}_{UU} + p_\gamma \boldsymbol{I}\right) \boldsymbol{f}_U \\ &= \lambda \boldsymbol{f}^T \boldsymbol{\Delta} \boldsymbol{f} + p_\gamma \boldsymbol{f}_U^T \boldsymbol{f}_U . \end{aligned} \tag{3}$$

□

Proof of Proposition 1

Proof. This proof is based on the observation that the sum of the 4 penalty terms

$$\|\boldsymbol{y}_U - \boldsymbol{\eta}_U^*\|_2^2 + \lambda_0 \|\boldsymbol{\beta}\|_2^2 + \sum_{\ell=1}^{q} \lambda_\ell \boldsymbol{f}^{(\ell)^T} \boldsymbol{\Delta}^{(\ell)} \boldsymbol{f}^{(\ell)} + \gamma \|\boldsymbol{y}_U\|_2^2$$

is a quadratic form with the vector of $n - m + p + nq$ unknown variables and a known, symmetric, positive semi-definite matrix. When this observation is combined with the use of labeled loss, optimal $\boldsymbol{y}_U$ and the $\boldsymbol{f}_U^{(\ell)}$ must be linear in $\boldsymbol{\beta}$, and the $\boldsymbol{f}_L^{(\ell)}$, so the existence of matrix $\boldsymbol{P}$ follows.

To see this, proceed as in the proof of Lemma 1 to get $\boldsymbol{y}_U = (1 - p_\gamma) \boldsymbol{\eta}_U$ analogous to (2) and profile out $\boldsymbol{y}_U$. The score for any $\boldsymbol{f}_U^{(\ell)}$ reduces to $\lambda_\ell \boldsymbol{\Delta}_{UL}^{(\ell)} \boldsymbol{f}_L^{(\ell)} + \lambda_\ell \boldsymbol{\Delta}_{UU}^{(\ell)} \boldsymbol{f}_U^{(\ell)} + p_\gamma \boldsymbol{\eta}_U^* = \vec{0}$, verifies

$$\left(\underset{\ell \le q}{\operatorname{diag}} \left(\lambda_\ell \boldsymbol{\Delta}_{UU}^{(\ell)} \right) + p_\gamma \boldsymbol{J} \otimes \boldsymbol{I} \right) \underset{\ell \le q}{\operatorname{vec}} \left(\widehat{\boldsymbol{f}}_U^{(\ell)} \right) = -\underset{\ell \le q}{\operatorname{vec}} \left(\lambda_\ell \boldsymbol{\Delta}_{UL}^{(\ell)} \widehat{\boldsymbol{f}}_L^{(\ell)} + \boldsymbol{X}_U \boldsymbol{\beta} \right),$$

and shows that each $\boldsymbol{f}_U^{(\ell)}$ is linear in $\boldsymbol{\beta}$ and $\left\{\boldsymbol{f}_L^{(\ell)}\right\}_{\ell=1}^q$. For any $\ell \in \{1, \cdots, q\}$,

$$\begin{aligned}\lambda_\ell \boldsymbol{f}^{(\ell)} \boldsymbol{\Delta}^{(\ell)} \boldsymbol{f}^{(\ell)} + p_\gamma \boldsymbol{\eta}_U^{*T} \boldsymbol{\eta}_U^* &= \lambda_\ell \boldsymbol{f}_L^{(\ell)T} \boldsymbol{\Delta}_{LL}^{(\ell)} \boldsymbol{f}_L^{(\ell)} - 2\lambda_\ell \boldsymbol{f}_U^{(\ell)T} \boldsymbol{\Delta}_{UL}^{(\ell)} \boldsymbol{f}_L^{(\ell)} \\ &\quad + \lambda_\ell \boldsymbol{f}_U^{(\ell)T} \boldsymbol{\Delta}_{UU}^{(\ell)} \boldsymbol{f}_U^{(\ell)} + p_\gamma \sum_{r=0}^{q} \boldsymbol{f}_U^{(r)T} \boldsymbol{f}_U^{(\ell)}\end{aligned}$$

with $\boldsymbol{f}_U^{(0)} = \boldsymbol{X}_U \boldsymbol{\beta}$. From this, reduction (3) generalizes to

$$\left\| \left(\operatorname*{diag}_{\ell \le q} \left(\lambda_\ell \boldsymbol{\Delta}_{LL}^{(\ell)} \right) \right)^{\frac{1}{2}} \operatorname*{vec}_{\ell \le q} \left(\boldsymbol{f}_L^{(\ell)} \right) \right\|_2^2 - \left\| \left(\operatorname*{diag}_{\ell \le q} \left(\lambda_\ell \boldsymbol{\Delta}_{UU}^{(\ell)} \right) + p_\gamma \boldsymbol{J} \otimes \boldsymbol{I} \right)^{-\frac{1}{2}} \operatorname*{vec}_{\ell \le q} \left(\lambda_\ell \boldsymbol{\Delta}_{UL}^{(\ell)} \boldsymbol{f}_L^{(\ell)} \right) \right\|_2^2,$$

which equals the penalty due to the lower-right $qm \times qm$ block of $\boldsymbol{P}$. The lower-left $qm \times p$ block of $\boldsymbol{P}$ is $p_\gamma \boldsymbol{X}_U$ stacked q times, while the upper-right $p \times qm$ block is its transpose. The upper-left $p \times p$ block of $\boldsymbol{P}$ will be $p_\gamma \boldsymbol{X}_U^T \boldsymbol{X}_U + \lambda_0 \boldsymbol{I}$. □

Proof of Proposition 2

Proof. The $\boldsymbol{y}_U$ terms in objective

$$\sum_{i \in L} \phi \left(y_i, \boldsymbol{T}_i \boldsymbol{\psi} \right) + \left\| \boldsymbol{y}_U - \boldsymbol{T}_U \boldsymbol{\psi} \right\|_2^2 + \boldsymbol{\psi}^T \boldsymbol{Q} \boldsymbol{\psi} + \gamma \left\| \boldsymbol{y}_U \right\|_2^2$$

have score $\boldsymbol{y}_U - \boldsymbol{T}_U \boldsymbol{\psi} + \gamma \boldsymbol{y}_U = \vec{0}$ which verifies $\boldsymbol{y}_U = (1 - p_\gamma) \boldsymbol{T}_U \boldsymbol{\psi}$. Vector $\boldsymbol{y}_U$ is then profiled out through the identity $\gamma \left\| \boldsymbol{y}_U \right\|_2^2 + \left\| \boldsymbol{T}_U \boldsymbol{\psi} - \boldsymbol{y}_U \right\|_2^2 = p_\gamma \left\| \boldsymbol{T}_U \boldsymbol{\psi} \right\|_2^2$. □

Proof of Theorem 1

Proof. Let $\widetilde{\boldsymbol{f}}_U = -\left(\boldsymbol{\Delta}_{UU} + \frac{p_\gamma}{\lambda} \boldsymbol{I} \right)^{-1} \boldsymbol{\Delta}_{UL} \widehat{\boldsymbol{f}}_{0L} = \sum_{i=1}^{n-m} \frac{\widehat{a}_{0i} \boldsymbol{o}_i}{d_i + p_\gamma}$. Consider error simplification

$$\begin{aligned}Q_1(\gamma) &= \mathbb{E}\left[\left\| \widehat{\boldsymbol{f}}_U - \boldsymbol{g}_U \right\|_2^2 \right] \\ &= \mathbb{E}\left[\left\| \widetilde{\boldsymbol{f}}_U - \boldsymbol{g}_U \right\|_2^2 \right] + 2\mathbb{E}\left[\left(\widetilde{\boldsymbol{f}}_U - \boldsymbol{g}_U \right)^T \left(\widehat{\boldsymbol{f}}_U - \widetilde{\boldsymbol{f}}_U \right) \right] + \mathbb{E}\left[\left\| \widetilde{\boldsymbol{f}}_U - \widehat{\boldsymbol{f}}_U \right\|_2^2 \right] \\ &= \mathbb{E}\left[\left\| \widetilde{\boldsymbol{f}}_U - \boldsymbol{g}_U \right\|_2^2 \right] + \sum_{i=1}^{n-m} \frac{\mathbb{E}\left[(\widehat{a}_{0i} - \widehat{a}_{\gamma i})^2 \right]}{(d_i + p_\gamma)^2} + 2 \sum_{i=1}^{n-m} \mathbb{E}\left[\left(\frac{\widehat{a}_{0i}}{d_i + p_\gamma} - b_i \right) \frac{(\widehat{a}_{0i} - \widehat{a}_{\gamma i})}{d_i + p_\gamma} \right] \\ &\le \mathbb{E}\left[\left\| \widetilde{\boldsymbol{f}}_U - \boldsymbol{g}_U \right\|_2^2 \right] + \epsilon^2 + \frac{2\epsilon}{\sqrt{n-m}} \sum_{i=1}^{n-m} \mathbb{E}\left[\left| \frac{\widehat{a}_{0i}}{d_i + p_\gamma} - b_i \right| \right] \\ &\le \mathbb{E}\left[\left\| \widetilde{\boldsymbol{f}}_U - \boldsymbol{g}_U \right\|_2^2 \right] + \frac{2M}{\sqrt{n-m}} \epsilon + \epsilon^2.\end{aligned}$$

From Assumption 2 of the main paper, let $r_i = \boldsymbol{o}_i^T \boldsymbol{\Sigma} \boldsymbol{o}_i$, so

$$Q_2\left(p_\gamma\right) = \mathbb{E}\left[\left\|\widetilde{\boldsymbol{f}}_U - \boldsymbol{g}_U\right\|_2^2\right] = \sum_{i=1}^{n-m} \mathbb{E}\left[\left(\frac{\widehat{a}_{0i}}{d_i+p_\gamma} - b_i\right)^2\right] = \sum_{i=1}^{n-m} \frac{r_i\sigma^2 + (a_i - b_i(d_i+p_\gamma))^2}{(d_i+p_\gamma)^2}. \tag{4}$$

The derivative of (4) with respect to p_γ evaluated at $p_\gamma = 0$ is

$$\begin{aligned} \left.\frac{\partial Q_2(p_\gamma)}{\partial p_\gamma}\right|_{p_\gamma=0} &= -2\sum_{i=1}^{n-m}\left.\left(\frac{r_i\sigma^2+(a_i-b_i(d_i+p_\gamma))^2}{(d_i+p_\gamma)^3} + \frac{(b_i(a_i-b_i(d_i+p_\gamma)))}{(d_i+p_\gamma)^2}\right)\right|_{p_\gamma=0} \\ &= -2\sum_{i=1}^{n-m}\frac{r_i\sigma^2+a_i^2-a_ib_id_i}{d_i^3} \end{aligned}$$

and is negative if

$$\sigma^2 > \left(\frac{\sum_{i=1}^{n-m}\left(a_ib_id_i - a_i^2\right)/d_i^3}{\sum_{i=1}^{n}\left\|\boldsymbol{\Sigma}^{1/2}\boldsymbol{o}_i\right\|_2^2/d_i^3}\right)_+ .$$

□

Proof of Theorem 2

Proof. If $\widetilde{\boldsymbol{w}}_j = \boldsymbol{R}^{1/2}\boldsymbol{w}_j$, then $\{\tau_j, \widetilde{\boldsymbol{w}}_j\}$ is an eigen-decomposition of $\boldsymbol{R}^{-1/2^T}\boldsymbol{T}_U^T\boldsymbol{T}_U\boldsymbol{R}^{-1/2} \succeq 0$ such that

$$\boldsymbol{w}_{j'}^T\boldsymbol{T}_U^T\boldsymbol{T}_U\boldsymbol{w}_j = \widetilde{\boldsymbol{w}}_{j'}^T\boldsymbol{R}^{-1/2^T}\boldsymbol{T}_U^T\boldsymbol{T}_U\boldsymbol{R}^{-1/2}\widetilde{\boldsymbol{w}}_j = \tau_j\widetilde{\boldsymbol{w}}_{j'}^T\widetilde{\boldsymbol{w}}_j = \tau_j\mathcal{I}_{\{j'=j\}}.$$

Let $\widetilde{\boldsymbol{c}} = \boldsymbol{R}^{1/2}\widehat{\boldsymbol{c}}$. Project $\widetilde{\boldsymbol{c}} = \sum_{j=1}^{\widetilde{p}}\widehat{a}_j\widetilde{\boldsymbol{w}}_j$, so $\widehat{\boldsymbol{c}} = \sum_{j=1}^{\widetilde{p}}\widehat{a}_j\boldsymbol{w}_j$ and $\widehat{\boldsymbol{\psi}} = \sum_{j=1}^{\widetilde{p}}\frac{\widehat{a}_j}{1+p_\gamma\tau_j}\boldsymbol{w}_j$. Let $r_j = \boldsymbol{w}_j^T\widetilde{\boldsymbol{\Sigma}}\boldsymbol{w}_j$. Thus,

$$\begin{aligned} Q_3\left(p_\gamma\right) &= \mathbb{E}\left[\left\|\boldsymbol{T}_U\left(\widehat{\boldsymbol{\psi}}_\gamma - \boldsymbol{\theta}\right)\right\|_2^2\right] \\ &= \sum_{j=1}^{\widetilde{p}}\tau_j\mathbb{E}\left[\left(\frac{\widehat{a}_j}{1+p_\gamma\tau_j} - q_j\right)^2\right] \\ &= \sum_{j=1}^{\widetilde{p}}\frac{\tau_j}{(1+p_\gamma\tau_j)^2}\left(r_j\sigma^2 + h_j\left(p_\gamma\right)\right), \end{aligned} \tag{5}$$

where $h_j\left(p_\gamma\right) = \left(\mu_j - q_j(1+p_\gamma\tau_j)\right)^2$. The derivative of (5) with respect to p_γ is

$$\left.\left(\frac{\partial Q_3(p_\gamma)}{\partial p_\gamma}\right)\right|_{p_\gamma=0} = \sum_{j=1}^{\widetilde{p}} -2\tau_j^2\left(r_j\sigma^2 + \mu_j^2 - \mu_jq_j\right) < 0$$

whenever σ^2 satisfies the given bound. □