Nonparametric Canonical Correlation Analysis
Supplementary Material

Tomer Michaeli TOMER.M@EE.TECHNION.AC.IL
Technion–Israel Institute of Technology, Haifa, Israel

Weiran Wang WEIRANWANG@TTIC.EDU
TTI-Chicago, Chicago, IL 60637, USA

Karen Livescu KLIVESCU@TTIC.EDU
TTI-Chicago, Chicago, IL 60637, USA

1. Proof of Lemma 3.1

Proof. Let the eigen-decomposition of the second-order moment of $\mathbb{E}[\mathbf{f}(X)|Y]$ be $\mathbb{E}[\mathbb{E}[\mathbf{f}(X)|Y]\mathbb{E}[\mathbf{f}(X)|Y]^\top] = \mathbf{A}\mathbf{D}\mathbf{A}^\top$ and define $U = \mathbf{A}^\top\mathbb{E}[\mathbf{f}(X)|Y]$ and $\tilde{\mathbf{g}}(Y) = \mathbf{A}^\top\mathbf{g}(Y)$. Then the objective in (3) can be written as $\mathbb{E}[\mathbf{f}(X)^\top\mathbf{g}(Y)] = \mathbb{E}[\mathbb{E}[\mathbf{f}(X)|Y]^\top\mathbf{g}(Y)] = \mathbb{E}[(\mathbf{A}^\top\mathbb{E}[\mathbf{f}(X)|Y])^\top(\mathbf{A}^\top\mathbf{g}(Y))] = \mathbb{E}[U^\top\tilde{\mathbf{g}}(Y)]$. Similarly, the constraint $\mathbf{I} = \mathbb{E}[\mathbf{g}(Y)\mathbf{g}(Y)^\top]$ can be expressed as $\mathbf{I} = \mathbf{A}^\top\mathbf{A} = \mathbb{E}[(\mathbf{A}^\top\mathbf{g}(Y))(\mathbf{A}^\top\mathbf{g}(Y))^\top] = \mathbb{E}[\tilde{\mathbf{g}}(Y)\tilde{\mathbf{g}}(Y)^\top]$. Therefore, the optimization problem (3) can be written in terms of $\tilde{\mathbf{g}}$ as

$$\max_{\tilde{\mathbf{g}}} \ \mathbb{E}\left[U^\top\tilde{\mathbf{g}}(Y)\right] \quad \text{s.t.} \quad \mathbb{E}\left[\tilde{\mathbf{g}}(Y)\tilde{\mathbf{g}}(Y)^\top\right] = \mathbf{I}. \tag{1}$$

Our objective is the sum of correlations in all L dimensions. Let us consider the correlation in the jth dimension. From the Cauchy-Schwartz inequality, we have

$$\mathbb{E}[U_j\tilde{g}_j(Y)] \leq \sqrt{\mathbb{E}\left[U_j^2\right]\mathbb{E}[\tilde{g}_j(Y)^2]} = \sqrt{\mathbb{E}\left[U_j^2\right]}$$

with equality if and only if $\tilde{g}_j(Y) = c_jU_j$ for some scalar c_j with probability 1. Note that choosing each $\tilde{g}_j(Y)$ to be proportional to U_j is valid, since the dimensions of U are uncorrelated (as $\mathbb{E}[UU^\top] = \mathbf{A}^\top\mathbb{E}\left[\mathbb{E}[\mathbf{f}(X)|Y]\mathbb{E}[\mathbf{f}(X)|Y]^T\right]\mathbf{A} = \mathbf{D}$). In order for each $\tilde{g}_j(Y)$ to have unit second order moment, we must have $c_j = 1/\sqrt{\mathbb{E}[U_j^2]} = 1/\sqrt{\mathbf{D}_{jj}}$. Therefore, $\tilde{\mathbf{g}}(Y) = \mathbf{D}^{-1/2}U$ so that $\mathbf{g}(Y) = \mathbf{A}\mathbf{D}^{-\frac{1}{2}}\mathbf{A}^\top U = (\mathbb{E}[\mathbb{E}[\mathbf{f}(X)|Y]\mathbb{E}[\mathbf{f}(X)|Y]^\top])^{-1/2}\mathbb{E}[\mathbf{f}(X)|Y]$, proving the lemma. $\square$