Inverse Regression for Ridge Recovery

Andrew Glaws[1], Paul Constantine[1], and Dennis Cook[2]
[1]Colorado School of Mines (contact:aglaws@mines.edu), [2]University of Minnesota

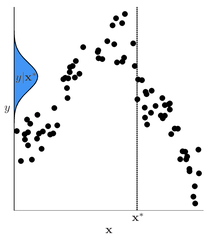
Sufficient Dimension Reduction

Regression problem:

Given: N iid samples drawn from $\pi(y, \mathbf{x})$

$$\left\{ \left[y_i \, , \, \mathbf{x}_i^T \right] \right\}, \quad y_i \in \mathbb{R}, \; \mathbf{x}_i \in \mathbb{R}^m$$

- often $\pi(y, \mathbf{x})$ is unknown



Goal: Estimate $y|\mathbf{x}$

Dimension reduction subspace (DRS):

For $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $n \leq m$, a dimension reduction subspace is

$$\mathcal{S}_{\text{DRS}} = \text{colspan}(\boldsymbol{A})$$

provided

$$y \perp\!\!\!\perp \mathbf{x} | \boldsymbol{A}^T \mathbf{x}.$$

If $\boldsymbol{A}$ generates a DRS, then

$$F_{y|\mathbf{x}}(Y) = F_{y|\boldsymbol{A}^T\mathbf{x}}(Y) \quad \text{for all } Y \in \mathbb{R}.$$

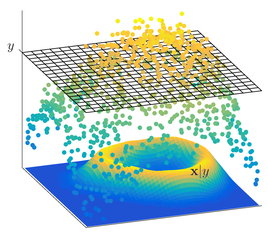
Inverse Regression: $\mathbf{x}|y$

Sliced Inverse Regression (SIR):

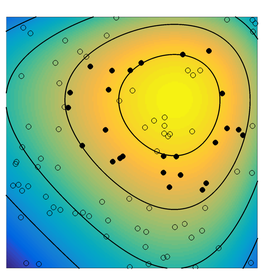
$$\text{Cov}\left[\mathbb{E}\left[\mathbf{x}|y\right]\right]$$

Sliced Average Variance Estimation (SAVE):

$$\mathbb{E}\left[\left(\boldsymbol{I} - \text{Cov}\left[\mathbf{x}|y\right]\right)^2\right]$$



SIR & SAVE Algorithms for Ridge Recovery



1. draw $\mathbf{x}_i$, $i = 1, \ldots, N$ according to $\rho(\mathbf{x})$
2. evaluate $y = f(\mathbf{x})$ for each $\mathbf{x}_i$ to obtain
$$\left\{ \left[y_i \, , \, \mathbf{x}_i^T \right] \right\}, \; i = 1, \ldots, N$$
3. partition the range into R slices,
$$J_r \text{ for } r = 1, \ldots, R$$
4. for SIR, compute
$$\hat{\boldsymbol{D}}_{\text{SIR}} = \frac{1}{N} \sum_{r=1}^{R} N_r \, \hat{\boldsymbol{\mu}}(r) \hat{\boldsymbol{\mu}}(r)^T$$
for SAVE, compute
$$\hat{\boldsymbol{D}}_{\text{SAVE}} = \frac{1}{N} \sum_{r=1}^{R} N_r \left[\boldsymbol{I} - \hat{\boldsymbol{\Sigma}}(r) \right]^2$$

Ridge Recovery

Deterministic function:

Given: deterministic function (e.g., computer simulation)

$$y = f(\mathbf{x}), \qquad y \in \mathbb{R}, \; \mathbf{x} \in \mathbb{R}^m$$

and input density $\rho(\mathbf{x})$

- $f(\mathbf{x})$ may be complicated or unknown

Goal: Study f

Ridge subspace:

For $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $n \leq m$, a ridge subspace is

$$\mathcal{S}_{\text{R}} = \text{colspan}(\boldsymbol{A})$$

provided

$$y = f(\mathbf{x}) = g(\boldsymbol{A}^T \mathbf{x})$$

for some $g : \mathbb{R}^n \to \mathbb{R}$.

Theorem

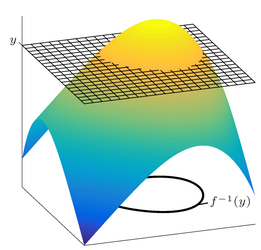
Suppose that $y = f(\mathbf{x})$ with $y \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^m$, and $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $n \leq m$. Then,

$$y \perp\!\!\!\perp \mathbf{x} | \boldsymbol{A}^T \mathbf{x} \qquad \Longleftrightarrow \qquad y = g(\boldsymbol{A}^T \mathbf{x}).$$

We define the central ridge subspace to be a subspace $\mathcal{S}_{f(\mathbf{x})}$ such that

- $\mathcal{S}_{f(\mathbf{x})}$ is a ridge subspace
- $\mathcal{S}_{f(\mathbf{x})} \subseteq \mathcal{S}_{\text{R}}$ for any ridge subspace $\mathcal{S}_{\text{R}}$

Inverse Image: $f^{-1}(y) = \{\mathbf{x} \in \mathbb{R}^m | y = f(\mathbf{x})\}$



Sliced Inverse Regression:

$$\boldsymbol{C}_{\text{SIR}} = \int \boldsymbol{\mu}(y) \boldsymbol{\mu}(y)^T \, d\gamma(y)$$

Sliced Average Variance Estimation:

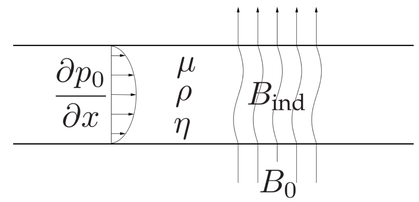
$$\boldsymbol{C}_{\text{SAVE}} = \int \left(\boldsymbol{I} - \boldsymbol{\Sigma}(y)\right)^2 \, d\gamma(y)$$

If $\rho(\mathbf{x})$ is elliptically symmetric, then

$$\text{colspan}(\boldsymbol{C}_{\text{SIR}}) \subseteq \mathcal{S}_{f(\mathbf{x})}$$

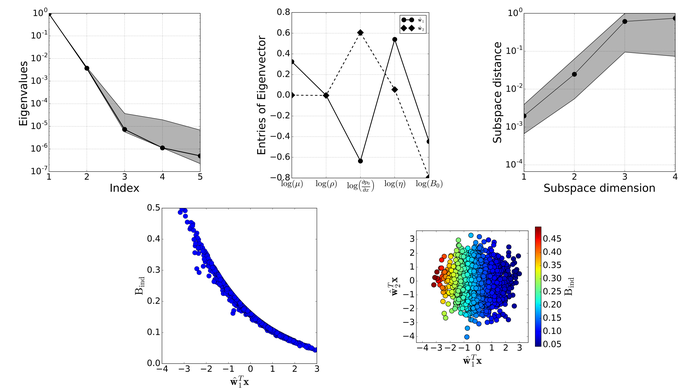
$$\text{colspan}(\boldsymbol{C}_{\text{SAVE}}) \subseteq \mathcal{S}_{f(\mathbf{x})}$$

Hartmann Problem

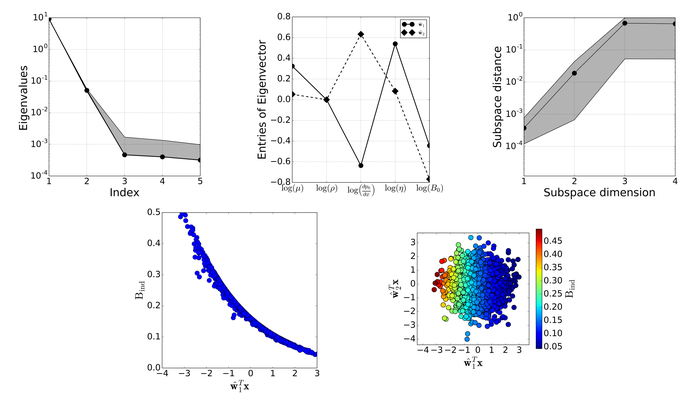


$$B_{\text{ind}} = f\left(\left[\log(\mu) \quad \log(\rho) \quad \log\left(\frac{\partial p_0}{\partial x}\right) \quad \log(\eta) \quad \log(B_0)\right]^T\right)$$

Sliced Inverse Regression:



Sliced Average Variance Estimation:



References

1 R. D. Cook, *Regression Graphics: Ideas for Studying Regression through Graphics*. John Wiley & Sons, Inc, New York (1998)
2 R. D. Cook, *SAVE: A method for dimension reduction and graphics in regression*. Communications in Statistics - Theory and Methods, 29 (2000), pp. 2109–2121
3 R. D. Cook and S. Weisberg, *Sliced inverse regression for dimension reduction: comment*. Journal of the American Statistical Association, 86 (1991), pp. 328–332
4 A. Glaws, P. G. Constantine, R. D. Cook, *Inverse regression for ridge recovery*. arXiv:1702.02227, (2017)
5 A. Glaws, P. G. Constantine, J. Shadid, and T. M. Wildey, *Dimension reduction in MHD power generation models: dimensional analysis and active subspaces*. arXiv:1609.01255, (2016)
6 K. C. Li, *Sliced inverse regression for dimension reduction*. Journal of the American Statistical Association, 86 (1991), pp. 316–327
7 A. Pinkus, *Ridge Functions*. Cambridge University Press, Cambridge (2015)