Active Subspaces

- We used sensitivity analysis to find influential parameters + then infer them.
  - We typically "fix" noninfluential parameters
  - We may transform $\vec{\theta}$ if not all are identifiable

- Instead, we can consider parameter space transformations.

ex) Consider

$$f(\vec{\theta}) = \exp\left(0.7\,\theta_1 + 0.3\,\theta_2\right), \quad \vec{\theta} \in \mathbb{R}^2$$

$$\frac{\partial f}{\partial \theta_1} = 0.7 \cdot f(\vec{\theta}), \quad \frac{\partial f}{\partial \theta_2} = 0.3\, f(\vec{\theta})$$

- Note: $\nabla_{\theta} f = \left[\frac{\partial f}{\partial \theta_1} \;\; \frac{\partial f}{\partial \theta_2}\right] = \left[0.7 f, \; 0.3 f\right] = \underset{\sim}{S}$

$$\Rightarrow F = \underset{\sim}{S}^T \underset{\sim}{S} \quad + \quad \operatorname{Cond}(F) = \infty$$

Linear Subspace Analysis

- Suppose we have

$$\vec{y} = \underset{\sim}{A}\vec{\theta}, \quad \vec{\theta} \in \mathbb{R}^p, \; \vec{y} \in \mathbb{R}^{N_y}, \; \underset{\sim}{A} \in \mathbb{R}^{N_y \times p}$$

Definition

The nonidentifiable subspace of the parameters, $NI(\vec{\theta})$, is

Definition

- The nonidentifiable subspace of the parameters, $NI(\vec{\theta})$, is where changes $\vec{\theta}$ have no effect on $\vec{y}$. (or non-unique effects)
  - This corresponds to $\text{null}(\underset{\sim}{A})$, i.e. where $\underset{\sim}{A}\vec{x} = \vec{0}$, $\forall \vec{x} \in \mathbb{R}^p$.

- The identifiable subspace, $I(\vec{\theta})$, is where $\vec{\theta}$ uniquely contributes to $\vec{y}$, + is the orthogonal complement of $\underset{\sim}{A}$.

Ex) $y_i = 0 \cdot \theta_1 + \theta_2 X_i$, $i = 1, 2, 3$ $\Rightarrow \vec{y} = \begin{bmatrix} 0 & x_1 \\ 0 & x_2 \\ 0 & x_3 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$

- So $\underset{\sim}{A} = \begin{bmatrix} 0 & x_1 \\ 0 & x_2 \\ 0 & x_3 \end{bmatrix} \Rightarrow \text{rank}(\underset{\sim}{A}) = 1$

$$\Rightarrow NI(\vec{\theta}) = \text{null}(\underset{\sim}{A}) = \left\{ c \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}, \ c \in \mathbb{R}$$

$$\Rightarrow I(\vec{\theta}) = \text{range}(\underset{\sim}{A}^T) = \mathcal{R}(\underset{\sim}{A}^T) = \left\{ c \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}, \ c \in \mathbb{R}$$

- Note: if $\underset{\sim}{A} = \underset{\sim}{X}$ from linear regression, we can show:

$$NI(\vec{\theta}) = \text{null}(\underset{\sim}{X}) \overset{?}{=} \text{null}(\underset{\sim}{X}^T \underset{\sim}{X}) = \text{null}(\underset{\sim}{F})$$

$\Rightarrow$ if $\underset{\sim}{F}$ is singular, $\vec{\theta}$ has some nonidentifiable components.

- If $\underset{\sim}{A} \in \mathbb{R}^{N_y \times P}$ + $N_y \geq P$, then $\vec{\theta}$ is only identifiable iff

$$\text{rank}(\underset{\sim}{A}) = P.$$

Computing Subspaces

- Recall that the SVD (singular-value decomposition) of $\underset{\sim}{A}$ is

$$\underset{\sim}{A} = \underset{\sim}{U}\,\underset{\sim}{\Sigma}\,\underset{\sim}{V}^T, \quad \underbrace{\underset{\sim}{U} \in \mathbb{R}^{N_y \times N_y}, \quad \underset{\sim}{V} \in \mathbb{R}^{P \times P}}_{\text{orthogonal}}$$

$$\& \quad \underset{\sim}{\Sigma} = \begin{bmatrix} \underset{\sim}{S} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} \end{bmatrix} \in \mathbb{R}^{N_y \times P}$$

$$\& \quad \underset{\sim}{S} = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_p) \in \mathbb{R}^{P \times P}, \text{ where } \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p \geq 0.$$

- The numerical rank of $\underset{\sim}{A}$ is the number of σ_i s.t. $\sigma_i \geq \varepsilon_{tol}$.

. When $\mathrm{rank}(\underset{\sim}{A}) = r < \min(N_y, P)$, $\underset{\sim}{A}$ is <u>rank deficient.</u>

- When $\underset{\sim}{A}$ is rank deficient, we can write (assume $\mathrm{rank}(\underset{\sim}{A}) = r$)

$$\Rightarrow \underset{\sim}{A} = \underset{\sim}{U}\,\underset{\sim}{\Sigma}\,\underset{\sim}{V}^T$$

where

$$\underset{\sim}{U} = [\underset{\sim}{U}_r \;\; \underset{\sim}{U}_{N_y - r}], \quad \underset{\sim}{U}_r \in \mathbb{R}^{N_y \times r}, \quad \underset{\sim}{U}_{N_y - r} \in \mathbb{R}^{N_y \times (N_y - r)}$$

$$\underset{\sim}{V} = [\underset{\sim}{V}_r \;\; \underset{\sim}{V}_{P-r}], \quad \underset{\sim}{V}_r \in \mathbb{R}^{P \times r}, \quad \underset{\sim}{V}_{P-r} \in \mathbb{R}^{P \times (P-r)}$$

$$\& \quad \underset{\sim}{\Sigma} = \begin{bmatrix} \underset{\sim}{S}_r & \underset{\sim}{S}_{r-p} \\ & \underset{\sim}{0} \end{bmatrix}, \quad \underset{\sim}{S}_r = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)$$

- Then we can approximate $\underset{\sim}{A}$ by

- Then we can approximate $\underset{\sim}{A}$ by

$$\underset{\sim}{A} \approx \underset{\sim}{U}_r \underset{\sim}{S}_r \underset{\sim}{V}_r^T$$

- Here, $\underset{\sim}{V}_r$ provides a basis $R(\underset{\sim}{A}^T)$, + $\underset{\sim}{V}_{p-r}$ provides a basis for null($\underset{\sim}{A}$).

* This is the truncated SVD *

- So for the linear model

$$\vec{y} = \underset{\sim}{A}\vec{\theta}$$

if rank($\underset{\sim}{A}$) $= r < \min(N_y, P)$, then

$$\Rightarrow \vec{y} \approx \underset{\sim}{A}_r \underset{\sim}{\theta} = \left(\underset{\sim}{U}_r \underset{\sim}{S}_r \underset{\sim}{V}_r^T\right)\vec{\theta}$$

where components of $\underset{\sim}{S}_r$ tell us about active subspaces in the model.