A Semiparametric Approach to Sufficient Dimension Reduction

MA, Yanyuan[1] ZHU, Liping[2]

[1]Texas A&M University

[2]Shanghai University of Finance and Economics

2012-10-05

An Outline

The Central Subspace

- The conditional distribution of Y depends to the covariates $\mathbf{x} = (X_1, \ldots, X_p)^\top$ only through $\boldsymbol{\beta}^\top\mathbf{x}$, where $\boldsymbol{\beta}$ is a $p \times d$ matrix and d is an unknown number.

$$\text{pr}(Y \leq y \mid \mathbf{x}) = \text{pr}(Y \leq y \mid \boldsymbol{\beta}^\top\mathbf{x}) \text{ for all } y \in \mathbb{R}.$$

- The above model can be written equivalently as

$$Y \perp\!\!\!\perp \mathbf{x} \mid \boldsymbol{\beta}^\top\mathbf{x}$$

Examples

- The linear model:
  $Y = \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x} + \varepsilon$;
  $Y \perp\!\!\!\perp \mathbf{x} \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}$ if $\varepsilon \perp\!\!\!\perp \mathbf{x}$.
- The index model:
  $Y = m(\boldsymbol{\alpha}_1^{\mathsf{T}}\mathbf{x}) + \sigma(\boldsymbol{\alpha}_2^{\mathsf{T}}\mathbf{x})\varepsilon$;
  $Y \perp\!\!\!\perp \mathbf{x} \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}$ when $\boldsymbol{\beta} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)$.
- The projection-pursuit model:
  $Y = m_1(\boldsymbol{\alpha}_1^{\mathsf{T}}\mathbf{x}) + \cdots + m_k(\boldsymbol{\alpha}_k^{\mathsf{T}}\mathbf{x}) + \varepsilon$;
  $Y \perp\!\!\!\perp \mathbf{x} \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}$ when $\boldsymbol{\beta} = (\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_k)$.

Examples

- The partially linear model:
  $Y = \boldsymbol{\alpha}_{\mathcal{L}}^{\mathsf{T}} \mathbf{x}_{\mathcal{L}} + m(\mathbf{x}_{\mathcal{N}}) + \varepsilon$;
  $Y \perp\!\!\!\perp \mathbf{x} \mid \boldsymbol{\beta}^{\mathsf{T}} \mathbf{x}$ when $\mathbf{x} = (\mathbf{x}_{\mathcal{L}}^{\mathsf{T}}, \mathbf{x}_{\mathcal{N}}^{\mathsf{T}})^{\mathsf{T}}$ and
  $$\boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\alpha}_{\mathcal{L}} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}.$$
- The partially linear single-index model:
  $Y = \boldsymbol{\alpha}_{\mathcal{L}}^{\mathsf{T}} \mathbf{x}_{\mathcal{L}} + m(\boldsymbol{\alpha}_{\mathcal{N}}^{\mathsf{T}} \mathbf{x}_{\mathcal{N}}) + \varepsilon$
  $Y \perp\!\!\!\perp \mathbf{x} \mid \boldsymbol{\beta}^{\mathsf{T}} \mathbf{x}$ when $\mathbf{x} = (\mathbf{x}_{\mathcal{L}}^{\mathsf{T}}, \mathbf{x}_{\mathcal{N}}^{\mathsf{T}})^{\mathsf{T}}$ and
  $$\boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\alpha}_{\mathcal{L}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\alpha}_{\mathcal{N}} \end{pmatrix}.$$

Examples

- The generalized linear model: $h\{E(Y \mid \mathbf{x})\} = \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}$; $Y \perp\!\!\!\perp \mathbf{x} \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}$.
  1. The logistic model: $\mathrm{logit}\{E(Y \mid \mathbf{x})\} = \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}$.
  2. The probit model: $E(Y \mid \mathbf{x}) = \Phi(\boldsymbol{\beta}^{\mathsf{T}}\mathbf{x})$.
  3. The Poisson model: $\log\{E(Y \mid \mathbf{x})\} = \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}$;
- The additive model: $Y = m_1(X_1) + \cdots + m_p(X_p) + \varepsilon$. $Y \perp\!\!\!\perp \mathbf{x} \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}$ when $\boldsymbol{\beta} = \mathbf{I}$.

Dimension Reduction Subspace

- $\boldsymbol{\beta}$ is not identifiable
  If $\boldsymbol{\beta}$ satisfies $Y \perp\!\!\!\perp \mathbf{x} \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}$, then $(\boldsymbol{\beta}, \boldsymbol{\beta}')$ satisfies as well.
  If $\boldsymbol{\beta}$ satisfies $Y \perp\!\!\!\perp \mathbf{x} \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}$, then $\boldsymbol{\beta}\mathbf{C}$ satisfies as well for any nonsingular $k \times k$ matrix $\mathbf{C}$. For example,

$$\boldsymbol{\beta} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } \boldsymbol{\beta} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

  span the same column space.
- The dimension reduction space is span$(\boldsymbol{\beta})$ if $Y \perp\!\!\!\perp \mathbf{x} \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}$.
- The central subspace: The smallest dimension reduction subspace $(\mathcal{S}_{Y|\mathbf{x}})$.

The Central Mean Subspace

- The conditional mean of Y relates to the covariates $\mathbf{x}$ only through $\boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}$.

$$E(Y \mid \mathbf{x}) = E(Y \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}).$$

- The central mean subspace $\mathcal{S}_{E(Y|\mathbf{x})}$: The column space of $\boldsymbol{\beta}$ if $\boldsymbol{\beta}$ has the smallest dimension.

Estimation of $\mathcal{S}_{Y|\mathbf{x}}$

- **SIR** Li 1991, JASA, Zhu and Fang 1996, AOS, Li and Dong 2009, AOS
- **SAVE** Cook and Weisberg 1991, JASA, Dong and Li 2010, BKA
- **DR** Li and Wang 2007, JASA
- **Fourier** Zhu and Zeng 2006, JASA
- **dMAVE** Xia 2007, AOS
- **DEE** Zhu, Wang, Zhu and Ferre 2009, BKA
- **CUME** Zhu, Zhu and Feng 2010, JASA

Estimation of $\mathcal{S}_{E(Y|\mathbf{x})}$

- **OLS** Li and Duan 1991, JASA
- **PHD** Li 1992, JASA and Cook and Li 2002, AOS
- **MAVE** Xia, Tong, Li and Zhu 2002, JRSSB

Distributional Assumptions

- For simplicity, we assume throughout that $E(\mathbf{x}) = \mathbf{0}$ and $\mathrm{cov}(\mathbf{x}) = \mathbf{I}$.
- The linearity condition (SIR, SAVE, DR, CUME, DEE, OLS, PHD)

$$E\left(\mathbf{x} \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\right) = \mathbf{P}\mathbf{x} = \boldsymbol{\beta}(\boldsymbol{\beta}^{\mathsf{T}}\boldsymbol{\beta})^{-1}\boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}$$

- The constant variance condition (SAVE, DR, CUME, DEE, PHD)

$$\mathrm{cov}\left(\mathbf{x} \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\right) = \mathbf{Q} = \mathbf{I}_p - \mathbf{P}$$

- The continuity of the covariates $\mathbf{x}$ (MAVE, dMAVE)

Our Ambition

- Derive a complete class of estimating equations for $\mathcal{S}_{Y|\mathbf{x}}$ and $\mathcal{S}_{E(Y|\mathbf{x})}$
- Eliminate all the conditions on the covariates

A Semiparametric Approach to Estimating $\mathcal{S}_{Y|\mathbf{x}}$

Using the geometric approach of Bickel, Klaassen, Ritov and Wellner (1993), we derive a complete class of estimating equations

$$\sum_{k=1}^{K} E\left[\left\{\mathbf{g}_k\left(Y, \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\right) - E\left(\mathbf{g}_k \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\right)\right\}\left\{\boldsymbol{\alpha}_k\left(\mathbf{x}\right) - E\left(\boldsymbol{\alpha}_k \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\right)\right\}\right] = \mathbf{0}.$$

Bickel, P. J., Klaassen, C. A. J., Ritov, Y. and Wellner, J. A. (1993) *Efficient and Adaptive Estimation for Semiparametric Models*. Baltimore: The Johns Hopkins University Press.

The Double Robustness Property

We let $K = 1$ to simplify the illustration of the double robustness property.

- When $K = 1$,

$$E\left[\{\mathbf{g}(Y, \beta^{\mathsf{T}}\mathbf{x}) - E(\mathbf{g} \mid \beta^{\mathsf{T}}\mathbf{x})\}\{\alpha(\mathbf{x}) - E(\alpha \mid \beta^{\mathsf{T}}\mathbf{x})\}\right] = \mathbf{0}.$$

- Mis-specify $E(\alpha \mid \beta^{\mathsf{T}}\mathbf{x})$ does not cause inconsistency.

$$E\left[\{\mathbf{g}(Y, \beta^{\mathsf{T}}\mathbf{x}) - E(\mathbf{g} \mid \beta^{\mathsf{T}}\mathbf{x})\}\{\alpha(\mathbf{x}) - \mathbf{h}(\beta^{\mathsf{T}}\mathbf{x})\}\right] = \mathbf{0}.$$

- Mis-specify $E(\mathbf{g} \mid \beta^{\mathsf{T}}\mathbf{x})$ does not cause inconsistency.

$$E\left[\{\mathbf{g}(Y, \beta^{\mathsf{T}}\mathbf{x}) - \mathbf{h}(\beta^{\mathsf{T}}\mathbf{x})\}\{\alpha(\mathbf{x}) - E(\alpha \mid \beta^{\mathsf{T}}\mathbf{x})\}\right] = \mathbf{0}.$$

SIR

- A complete class of estimating equations

$$\sum_{k=1}^{K} E\left[\left\{\mathbf{g}_k\left(Y,\beta^{\mathsf{T}}\mathbf{x}\right)-E\left(\mathbf{g}_k \mid \beta^{\mathsf{T}}\mathbf{x}\right)\right\}\left\{\alpha_k\left(\mathbf{x}\right)-E\left(\alpha_k \mid \beta^{\mathsf{T}}\mathbf{x}\right)\right\}\right]=\mathbf{0}.$$

- SIR sets $\mathbf{g}(Y,\beta^{\mathsf{T}}\mathbf{x}) = E(\mathbf{x} \mid Y)$ and $\alpha(\mathbf{x}) = \mathbf{x}^{\mathsf{T}}$.
- SIR assumes that $E(\mathbf{x} \mid \beta^{\mathsf{T}}\mathbf{x}) = \mathbf{P}\mathbf{x}$.
- It mis-specifies $E\left(\mathbf{g} \mid \beta^{\mathsf{T}}\mathbf{x}\right) = \mathbf{0}$.
- $\mathrm{cov}\{E(\mathbf{x} \mid Y)\}\mathbf{Q} = \mathbf{0}$, where $\mathbf{Q} = \mathbf{I} - \mathbf{P} = \mathbf{I} - \beta(\beta^{\mathsf{T}}\beta)^{-1}\beta^{\mathsf{T}}$.
- Improve SIR by estimating $E\left(\mathbf{g}_k \mid \beta^{\mathsf{T}}\mathbf{x}\right)$ and $E(\mathbf{x} \mid \beta^{\mathsf{T}}\mathbf{x})$ nonparametrically.

SAVE

- A complete class of estimating equations

$$\sum_{k=1}^{K} E\left[\left\{\mathbf{g}_k\left(Y, \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\right) - E\left(\mathbf{g}_k \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\right)\right\}\left\{\boldsymbol{\alpha}_k\left(\mathbf{x}\right) - E\left(\boldsymbol{\alpha}_k \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\right)\right\}\right] = \mathbf{0}.$$

- SAVE sets $\mathbf{g}_1(Y, \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}) = \mathbf{I}_p - \mathrm{cov}(\mathbf{x} \mid Y)$, $\mathbf{g}_2(Y, \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}) = \mathbf{g}_1 E(\mathbf{x} \mid Y)$, $\boldsymbol{\alpha}_1(\mathbf{x}) = -\mathbf{x}\{\mathbf{x} - E(\mathbf{x} \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x})\}^{\mathsf{T}}$, and $\boldsymbol{\alpha}_2(\mathbf{x}) = \mathbf{x}^{\mathsf{T}}$.
- SAVE assumes both the linearity and the constant variance conditions to ensure that $E(\boldsymbol{\alpha}_k \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x})$ have specific parametric forms. It mis-specifies $E\left(\mathbf{g}_k \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\right) = \mathbf{0}$.
- $E\left\{\mathbf{I} - \mathrm{cov}(\mathbf{x} \mid Y)\right\}^2 \mathbf{Q} = \mathbf{0}$.
- Improve SAVE by estimating $E(\boldsymbol{\alpha}_k \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x})$ nonparamerically.

DR

- A complete class of estimating equations

$$\sum_{k=1}^{K} E\left[\left\{\mathbf{g}_k\left(Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\right) - E\left(\mathbf{g}_k \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\right)\right\}\left\{\boldsymbol{\alpha}_k\left(\mathbf{x}\right) - E\left(\boldsymbol{\alpha}_k \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\right)\right\}\right] = \mathbf{0}.$$

- DR sets $\mathbf{g}_1(Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}) = \mathbf{I}_p - E(\mathbf{x}\mathbf{x}^{\mathrm{T}} \mid Y)$,
  $\mathbf{g}_2(Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}) = E\left\{E(\mathbf{x} \mid Y)E(\mathbf{x}^{\mathrm{T}} \mid Y)\right\} E(\mathbf{x} \mid Y)$
  $\mathbf{g}_3(Y, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}) = E\left\{E(\mathbf{x}^{\mathrm{T}} \mid Y)E(\mathbf{x} \mid Y)\right\} E(\mathbf{x} \mid Y)$
  $\boldsymbol{\alpha}_1(\mathbf{x}) = -\mathbf{x}\{\mathbf{x} - E(\mathbf{x} \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})\}^{\mathrm{T}}$, $\boldsymbol{\alpha}_2(\mathbf{x}) = \boldsymbol{\alpha}_3(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}$.
- DR assumes that $E(\boldsymbol{\alpha}_k \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})$ have specific parametric forms using the linearity and the constant variance conditions. It mis-specifies $E\left(\mathbf{g}_k \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}\right) = \mathbf{0}$.
- Improve DR by estimating $E(\boldsymbol{\alpha}_k \mid \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x})$ nonparamerically.

A Semiparametric Approach to Estimating $\mathcal{S}_{E(Y|\mathbf{x})}$

A complete class of estimating equations

$$E\left[\{Y - E(Y \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x})\}\left\{\boldsymbol{\alpha}(\mathbf{x}) - E\left(\boldsymbol{\alpha} \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\right)\right\}\right] = \mathbf{0}.$$

Compare to the estimating equations for estimating $\mathcal{S}_{Y|\mathbf{x}}$,

$$E\left[\left\{\mathbf{g}\left(Y, \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\right) - E\left(\mathbf{g} \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\right)\right\}\left\{\boldsymbol{\alpha}\left(\mathbf{x}\right) - E\left(\boldsymbol{\alpha} \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\right)\right\}\right] = \mathbf{0}.$$

The Double Robustness Property

- Mis-specify $E(\boldsymbol{\alpha} \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x})$ does not cause inconsistency.

$$E\left[\{Y - E(Y \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x})\}\{\boldsymbol{\alpha}(\mathbf{x}) - \mathbf{h}(\boldsymbol{\beta}^{\mathsf{T}}\mathbf{x})\}\right] = \mathbf{0}.$$

- Mis-specify $E(Y \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x})$ does not cause inconsistency.

$$E\left[\{Y - \mathbf{h}(\boldsymbol{\beta}^{\mathsf{T}}\mathbf{x})\}\{\boldsymbol{\alpha}(\mathbf{x}) - E(\boldsymbol{\alpha} \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x})\}\right] = \mathbf{0}.$$

OLS

- A complete class of estimating equations

$$E\left[\{Y - E(Y \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x})\}\left\{\boldsymbol{\alpha}(\mathbf{x}) - E\left(\boldsymbol{\alpha} \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\right)\right\}\right] = \mathbf{0}.$$

- OLS sets $\boldsymbol{\alpha}(\mathbf{x}) = \mathbf{x}$.
- OLS assumes $E\{\boldsymbol{\alpha}(\mathbf{x}) \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\} = \mathbf{P}\mathbf{x}$ and mis-specifies $E\left(Y \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\right) = \mathbf{0}$ to obtain that $E(\mathbf{x}Y)\mathbf{Q} = \mathbf{0}$.
- Improve OLS by estimating $E(Y \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x})$ and $E\{\boldsymbol{\alpha}(\mathbf{x}) \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\}$ nonparametrically.

PHD

- A complete class of estimating equations

$$E\left[\{Y - E(Y \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x})\}\left\{\boldsymbol{\alpha}(\mathbf{x}) - E\left(\boldsymbol{\alpha} \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\right)\right\}\right] = \mathbf{0}.$$

- PHD sets $\boldsymbol{\alpha}(\mathbf{x}) = \mathbf{x}\mathbf{x}^{\mathsf{T}}$ and mis-specifies $E\left(Y \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\right) = E(Y)$.
- PHD assumes that both the linearity and the constant variance conditions, which insures that $\boldsymbol{\alpha}(\mathbf{x}) - E\left\{\boldsymbol{\alpha}(\mathbf{x}) \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}\right\} = \mathbf{x}\mathbf{x}^{\mathsf{T}} - \mathbf{Q} - \mathbf{P}\mathbf{x}\mathbf{x}^{\mathsf{T}}\mathbf{P}$.
- $\mathbf{Q}E\{Y - E(Y)\}\mathbf{x}\mathbf{x}^{\mathsf{T}}\mathbf{Q} = \mathbf{0}$.
- Improve PHD by $E(Y \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x})$ and $E(\mathbf{x}\mathbf{x}^{\mathsf{T}} \mid \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x})$ nonparametrically.

Simulated Examples

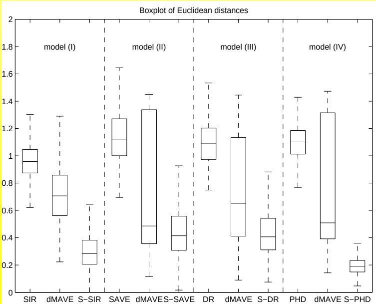
- We set $p = 6$, $d = 2$ and $n = 200$.
- We use the following models.

$$\begin{aligned}
\text{model 1}: &\quad Y = (\mathbf{x}^\top\beta_1)\,/\left\{0.5 + (\mathbf{x}^\top\beta_2 + 1.5)^2\right\} + 0.5\varepsilon;\\
\text{model 2}: &\quad Y = (\mathbf{x}^\top\beta_1)^2 + 2\mid \mathbf{x}^\top\beta_2 \mid + 0.1 \mid \mathbf{x}^\top\beta_2 \mid \varepsilon;\\
\text{model 3}: &\quad Y = \exp(\mathbf{x}^\top\beta_1) + 2(\mathbf{x}^\top\beta_2 + 1)^2 + \mid \mathbf{x}^\top\beta_1 \mid \varepsilon;\\
\text{model 4}: &\quad Y = (\mathbf{x}^\top\beta_1)^2 + (\mathbf{x}^\top\beta_2)^2 + 0.5\varepsilon,
\end{aligned}$$

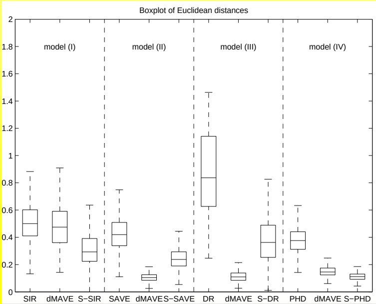
SIR and Semi-SIR

$$
\begin{aligned}
E[E(\mathbf{x} \mid Y)\{\mathbf{x} - E(\mathbf{x} \mid \beta^{\mathsf{T}}\mathbf{x})\}] &= \mathbf{0}; \\
E[E(\mathbf{x} \mid Y)\{\mathbf{x} - \widehat{E}(\mathbf{x} \mid \beta^{\mathsf{T}}\mathbf{x})\}] &= \mathbf{0}.
\end{aligned}
$$

Distributional assumptions are violated



Distributional assumptions are satisfied



Summary

1. We derive a complete class of estimating equations for estimating $\mathcal{S}_{Y|\mathbf{x}}$ and $\mathcal{S}_{E(Y|\mathbf{x})}$.
2. The semiparametric approaches require no assumptions on the covariates.

Some Questions

1. In the complete class, which is the optimal one? How to define the optimality? Is an efficient estimate of $\mathcal{S}_{Y|\mathbf{x}}$ available?
2. Our simulations show that, even when both the linearity and the constant variance conditions are satisfied, the semiparametric approaches still outperform the classical methods. Is such an improvements real or marginal?
3. How to estimate the dimension of $\mathcal{S}_{Y|\mathbf{x}}$ and $\mathcal{S}_{E(Y|\mathbf{x})}$ within the semiparametric framework?

The End