Dimension reduction by propensity score

Can we use a propensity score instead of covariates?

Consider the regression model[1]

$$y = \beta_0 + \beta_d d + \beta_x x + \epsilon \tag{1}$$

where y is the outcome or dependent variable, d is a dummy variable representing group/class membership, and x is a single covariate. Here d and x are the independent or predictor variables. To simplify the math we will assume that x is centered.

We'll use the Frisch-Waugh-Lovell (FWL) Theorem to estimate β_d.

The FWL theorem states that any predictor's regression coefficient in a multivariate model is equivalent to the regression coefficient estimated from a bivariate model in which the residualised outcome is regressed on the residualized component of the predictor; where the residuals are taken from models regressing the outcome **and** the predictor on all other predictors in the multivariate regression (separately).

First residual - the propensity score:

The propensity score predicts the dummy variable using the covariates, i.e the regression:

$$d = \beta_x^{(d)} x + \epsilon^{(d)}$$

where $\beta_x^{(d)} x = x\frac{\text{cov}(x,d)}{\text{var}(x)}$. We define the propensity score as $P(x) = x\frac{\text{cov}(x,d)}{\text{var}(x)}$.

Thus $d - P(x)$ is the first leg of the FWL Theorem applied to Equation 1.

[1] this can be generalized to multiple covariates.

Second residual:

The second leg of the FWL Theorem starts with the regression:

$$y = \beta_x^{(y)} x + \epsilon^{(y)} \tag{2}$$

where $\beta_x^{(y)} = \frac{\text{cov}(x,y)}{\text{var}(x)}$, and the second leg of the FWL Theorem is $y - x\frac{\text{cov}(x,y)}{\text{var}(x)}$.

Can we regress on $P(x)$ instead of x in Equation 2?

For our second leg, instead of Equation 2 can we use the following regression?

$$y = \beta_{P(x)}^{(y)} P(x) + \epsilon^{(y)} \tag{3}$$

We can write Equation 3 as $y = P(x)\frac{\text{cov}(P(x),y)}{\text{var}(P(x))} + \epsilon^{(y)}$, and then note that:

- $P(x)\frac{\text{cov}(P(x),y)}{\text{var}(P(x))} = P(x)\left(P(x)^\top P(x)\right)^{-1}\text{cov(P(x),y)}$, and

$$P(x)\left(P(x)^\top P(x)\right)^{-1}\text{cov(P(x),y)} = x\frac{\text{cov}(d,x)}{\text{var}(x)} \times \frac{\text{var}(x)}{\left(\text{cov}(d,x)\right)^2} \times \frac{\text{cov}(d,x)}{\text{var}(x)}\text{cov}(x,y)$$

- since $P(x)^\top P(x) = \frac{\text{cov}(d,x)}{\text{var}(x)} x^\top x \frac{\text{cov}(d,x)}{\text{var}(x)} = \frac{\left(\text{cov}(d,x)\right)^2}{\text{var}(x)}$, and
- $\text{cov}(P(x),y) = \text{cov}(\frac{\text{cov}(d,x)}{\text{var}(x)}x, y) = \frac{\text{cov}(d,x)}{\text{var}(x)}\text{cov}(x,y)$
- so $P(x)\frac{\text{cov}(P(x),y)}{\text{var}(P(x))} = x\frac{\text{cov}(x,y)}{\text{var}(x)}$ and they are exchangeable.

It follows that in the regression $y = \beta_0 + \beta_d d + \beta_x x + \epsilon$, with the propensity score $P(x) = \frac{\text{cov}(d,x)}{\text{var}(x)}x$, then per the Frisch-Waugh theorem the coefficient β_d is equal to the coefficient from regressing

$$y - \frac{\text{cov}(P(x),y)}{\text{var}(P(x))}P(x) \text{ on } d - P(x)$$

With multiple covariates:

- For a covariate matrix X the propensity score is written as $P(X) = X\frac{X^\top d}{X^\top X} = X(X^\top X)^{-1}X^\top d$
- For $P(x)\frac{\text{cov}(P(x),y)}{\text{var}(P(x))}$
  - $\text{var}(P(x)) = P(x)^\top P(x)$ which is

    $$\left(X(X^\top X)^{-1}X^\top d\right)^\top \times X(XX^\top)^{-1}X^\top d = d^\top X(X^\top X)^{-1}X^\top \times X(X^\top X)^{-1}X^\top d$$

    which simplifies to $d^\top X(XX^\top)^{-1}X^\top d$
- $\left(P(x)^\top P(x)\right)^{-1}$ is $\left(d^\top X(XX^\top)^{-1}X^\top d\right)^{-1}$ which is $\left(X^\top d\right)^{-1}(XX^\top)(d^\top X)^{-1}$
- $\text{cov}(P(x), y)$ is $P(x)^\top y$ which is $\left(X(X^\top X)^{-1}X^\top d\right)^\top y = d^\top X(XX^\top)^{-1}X^\top y$, so
- $\frac{\text{cov}(P(x),y)}{\text{var}(P(x))}$ is $\left(X^\top d\right)^{-1}(XX^\top)(d^\top X)^{-1}d^\top X(XX^\top)^{-1}X^\top y = \left(X^\top d\right)^{-1}X^\top y$, and finally
- $P(x)\frac{\text{cov}(P(x),y)}{\text{var}(P(x))} = X(X^\top X)^{-1}X^\top d\left(X^\top d\right)^{-1}X^\top y$ which is $X(X^\top X)^{-1}X^\top y = X\frac{\text{cov}(X,y)}{\text{var}(X)}$