Motivation for CPOP: one patient cohort, two gene expression data

$$Z_1 \hat{\beta}_1 \approx Z_2 \hat{\beta}_2$$
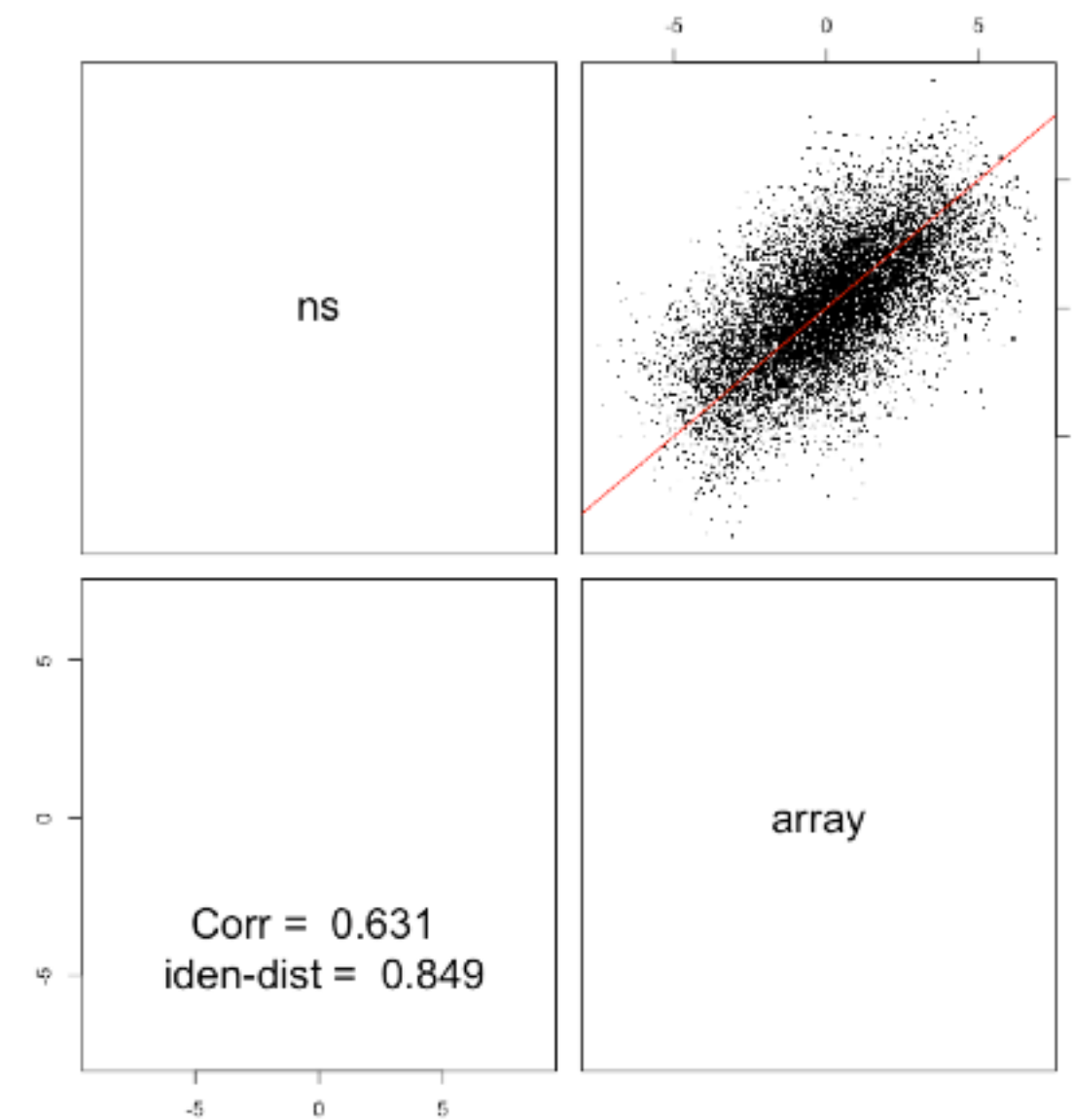
means

$$Z_1 \approx Z_2 \quad \text{column-wise}$$

$$\hat{\beta}_1 \approx \hat{\beta}_2 \quad \text{element-wise}$$

CPOP framework: weighted variable selection

1. Perform a **weighted Lasso** by placing higher weights on features closer to the identity line

2. Perform a **Ridge regression** and only retain those features with coefficients similar to each other

$$Z_1 \approx Z_2$$

$$\hat{\beta}_1 \approx \hat{\beta}_2$$