$$S_0 = \{ j ; f_j^0(\cdot) \not\equiv 0 \} \subseteq \{1 \to p\}$$

$$\hat{S} = \{ j ; \hat{f}_j(\cdot) \not\equiv 0 \}$$

$$\hat{S} \overset{?}{\geq} S_0$$

<u>fulklore "theorem"</u>: it works if

$$\sqrt{s_0 \frac{\log(\text{dimensionality})}{n^{4/5}}} \text{ is } \sqrt{\text{small}}$$

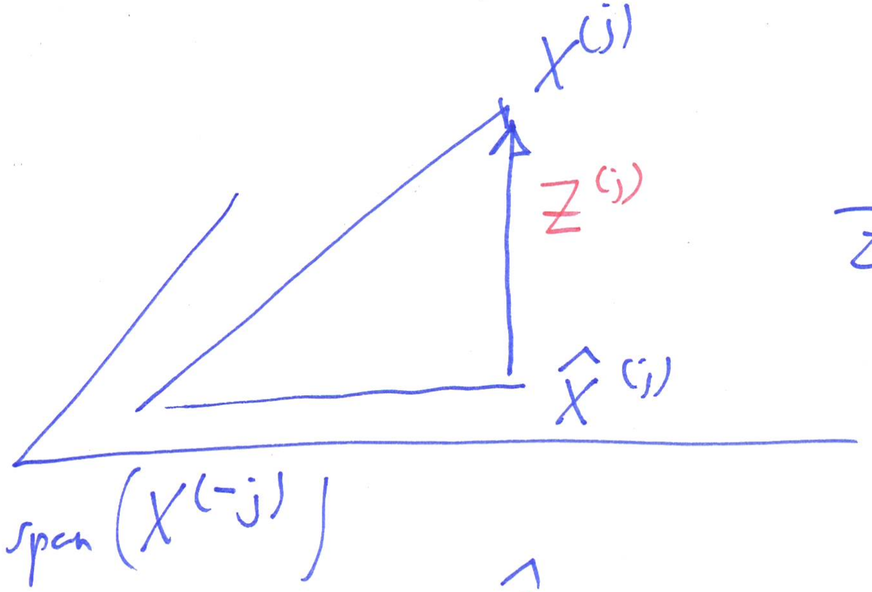
OK ... why not interaction modeling

$$Y_i = \mu + \sum_{j=1}^{p} f_j(X_i^{(j)}) + \underbrace{\sum_{k \neq l} g_{kl}(X_i^{(k)}, X_i^{(l)})}_{\mathcal{O}(p^2)} + \varepsilon_i$$

$$\rightsquigarrow \dim = \mathcal{O}(p^2 \cdot n)$$

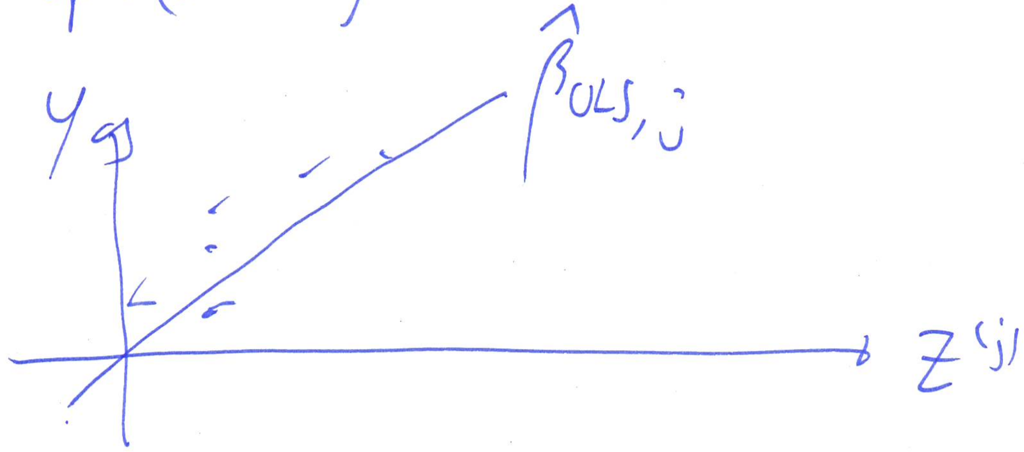
$$\underline{\log(\dim) = \mathcal{O}(\log(p))} \quad \text{if } p \gg n$$

first interaction modeling is computationally doable
second " " : more clever techniques



$Z^{(j)} \perp\!\!\!\perp \hat{X}^{(j)}$

$Z^{(j)} \perp\!\!\!\perp X^{(-j)}$



$$Y = X\beta^0 + \varepsilon$$

$$= \sum_{k=1}^{p} \beta_k^0 X^{(k)} + \varepsilon$$

$$= \beta_j^0 X^{(j)} + \sum_{k \neq j} \beta_k^0 X^{(k)} + \varepsilon$$

$$\frac{Y^T Z^{(j)}}{(X^{(j)})^T Z^{(j)}} = \beta_j^0 + \sum_{k \neq j} \frac{(X^{(k)})^T Z^{(j)}}{(X^{(j)})^T Z^{(j)}} \beta_k^0 + \frac{\varepsilon^T Z^{(j)}}{(X^{(j)})^T Z^{(j)}}$$