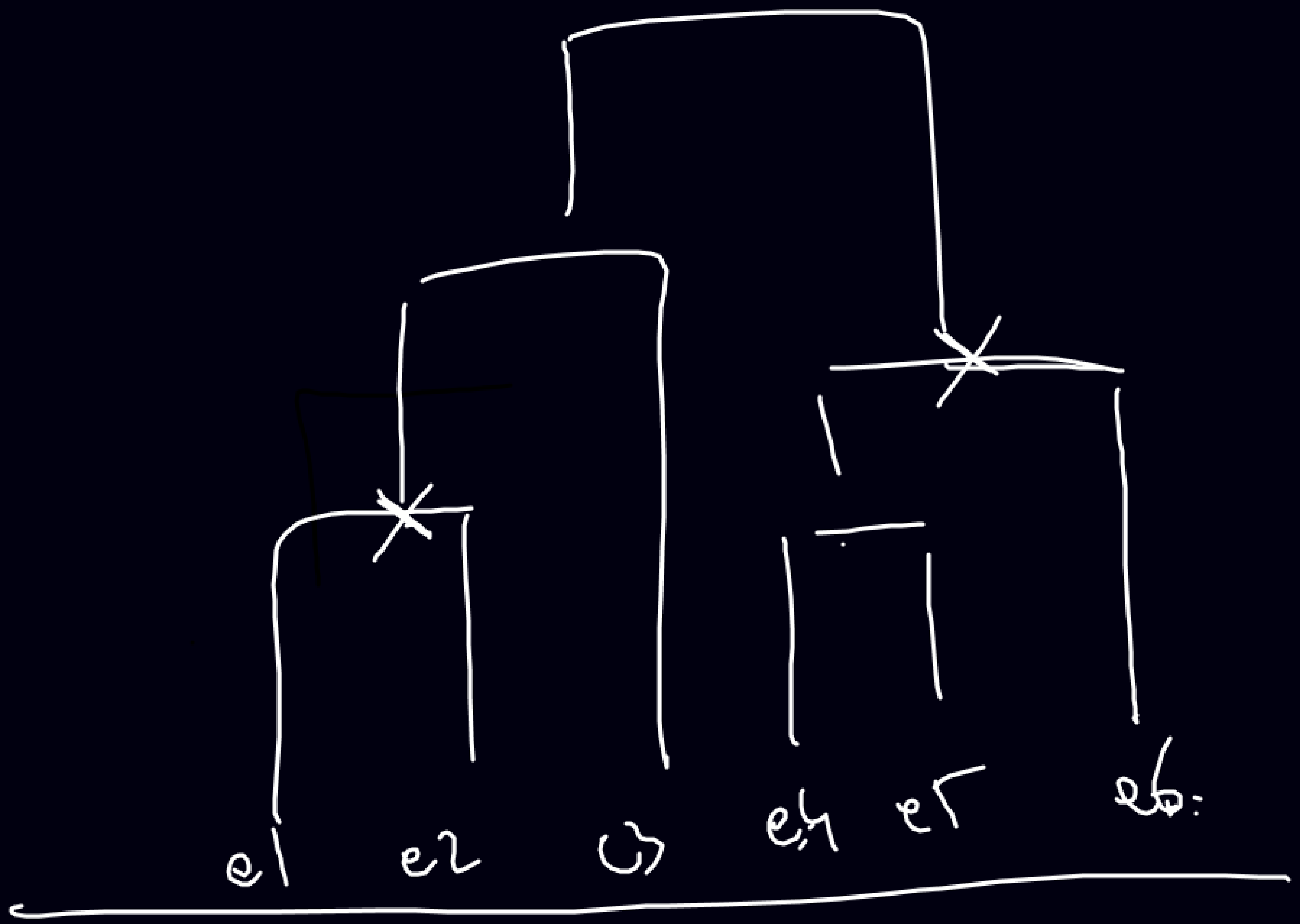
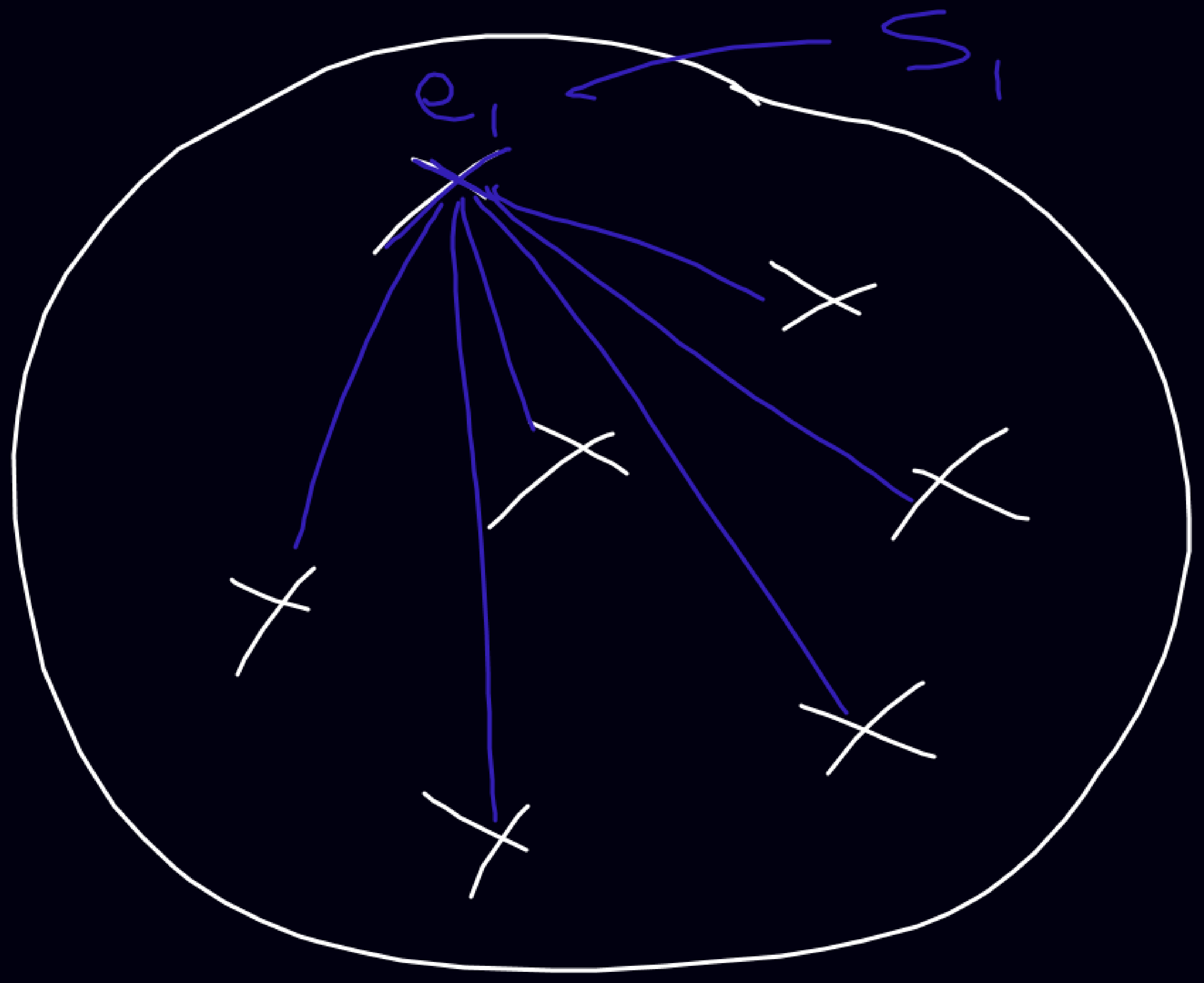

e1
e2
e3
e4
e5
e6



e_1
S_1

$$\beta = \begin{pmatrix} \mu \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \alpha_1 \\ \alpha_2 \\ \beta_6 \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \\ \beta_8 \end{pmatrix} \quad \vdots$$

$$LH = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$L = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

$$\left.\begin{array}{l}\text{Ridge}\\ \text{Lasso}\end{array}\right\}$$ penalized version of the linear regression.

optimization p^b in classical linear regression is:

$$\min_{\beta \in \mathbb{R}^{p+1}} \left\| Y - \left(\mu_0 + \sum_{i=1}^{p} \beta_i X^i\right) \right\|_2^2$$

Now: $p >> n$

$$\mathrm{rk}(X) \leq n < p \Rightarrow {}^{t}XX$$ not invertible

$$X = \left(1 \mid X^1 \mid X^2 \mid X^3 \mid \cdots \mid X^p \right)$$

We have $p >> n$

↳ the information inside the p variables is repeated.

⇒ my model: $Y = f(X^1, \dots, X^p) + \varepsilon$

↑ linear f^{cd}

In fact: $Y = g(\{X^i, i \in I \text{ with } I \subset \{1, \dots, p\}\}) + \varepsilon$

↳ notion of sparsity.

Ridge:

optimization problem:

$$\min_{\substack{\beta \in \mathbb{R}^{p+1} \\ \mu^2 + \sum_{i=1}^{p} \beta_i^2 \leq s}} \left\| Y - \left(\mu + \sum_{i=1}^{p} \beta_i X^i\right) \right\|^2$$

with $s \geq 0$!

$$\Leftrightarrow \min_{\beta \in \mathbb{R}^{p+1}} \left(\left\| Y - \left(\mu + \sum_{i=1}^{p} \beta_i X^i\right) \right\|^2 + \lambda \sum_{i=0}^{p} \beta_i^2 \right)$$

with $\beta_0 = \mu$

Rk: $\lambda \to 0$
we find Ordinary least Square

$\lambda \to +\infty$ all β_i closed to 0!

to solve the p^b:

$$-2{}^tX(Y - X\beta) + 2\lambda\beta = 0$$

$$X = \begin{pmatrix} 1 \,\Big|\, X^1 \,\Big|\, \cdots \,\Big|\, X^p \end{pmatrix} \quad \beta = \begin{pmatrix} \mu \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}$$

$$\hat{\beta}_\lambda^{ridge} = \left({}^tXX + \underbrace{\lambda I_p}\right)^{-1} {}^tXY$$

P^b: How to choose λ?

↳ by cross-validation

$$E\left[\hat{\beta}_\lambda^{ridge}\right] = \left(I_p + \lambda R^{-1}\right)^{-1}\beta$$

with $R = {}^tXX$

$$\hat{\beta}_{\lambda}^{ridge} = \left(I_p + \lambda R^{-1} \right)^{-1} \hat{\beta}^{LS}$$

Lasso:

$$\min_{\beta \in \mathbb{R}^p} \left(\|Y - X\beta\|^2 + \lambda \sum_{i=0}^{p} |\beta_i| \right)$$

↳ no analytic expression of the solution!