POPULAT.	SAMPLE
$y = \beta' x + \epsilon$	$y_i = \hat{\beta}' x_i + \hat{\epsilon}_i$
$E(y^2) = E(\beta' x)^2 + E(\epsilon^2)$	$\sum y_i^2 = \sum (\beta' x_i)^2 + \sum \hat{\epsilon}_i^2$
$MSE_{POP} = E\epsilon^2$	$MSE_{SAMPLE} = \sum \hat{\epsilon}_i^2$
$E(y - \beta' x)^2$	
$R^2_{POP} = 1 - \frac{E\epsilon^2}{Ey^2}$	$R^2_{SAM} = 1 - \frac{\sum \hat{\epsilon}_i^2}{\sum y_i^2}$
$\in [0,1]$	$\in [0,1]$

$X \sim N(0, I_p)$ $Y \sim N(0,1)$

If $p = n$ (100 100) → overfitting

$R^2_{sample} = 1$ ✓

$p = \frac{n}{2}$ $\quad R^2_{sample} \cong 0{,}5$

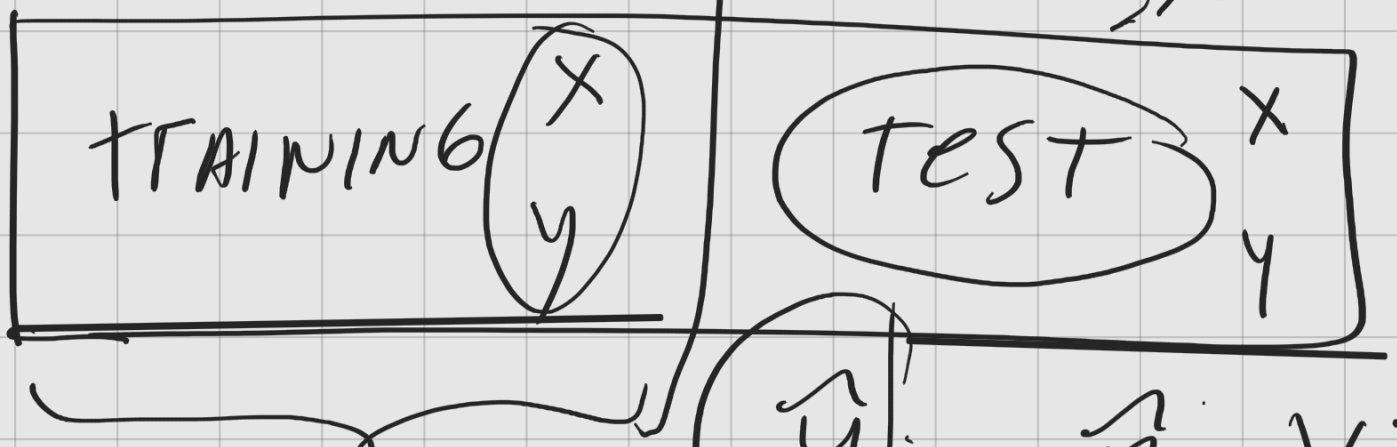
$p = \frac{n}{20}$ $\quad R^2_{sample} \approx 0{,}05$

$$MSE_{Adjusted} = \frac{n}{n-p} \sum_n \hat{\epsilon}_i^2$$

$$R^2_{Adjusted} = \left(1 - \frac{n}{n-p}\right) \frac{\sum_n \hat{\epsilon}_i^2}{\sum_n y_i^2}$$

* SAMPLE SPLITTING

0,75 | 0,25



$y_{tr} = \beta' X_{tr} + \epsilon_{tr}$

$\hat{y} = \hat{\beta}_{TRAIN} X_{TEST}$

B

y_{TEST} VS $\hat{y}_{TEST}$

① Sample SPLITTING

② Training data $\Rightarrow \hat{f}(X)$

$$\hat{f}(X) = \beta' X$$

③ $$MSE_{test} = \frac{1}{n} \sum_{k \in V} \left(y_k - \hat{f}(x_k) \right)^2$$

$\hat{y}$

$V \Rightarrow$ ~~Test~~ Test Sample

MSE_{test} ⇒ OUT OF SAMPLE MSE

$$R^2_{test} = 1 - \frac{MSE_{test}}{\frac{1}{m}\sum_{k \in N} y_k^2}$$

TRAIN ← 0,5 n | 0,5 n → TEST

X_1 y_1 | X_2 y_2

$y_1 \sim X_1$

$\hat{\beta}_1$

$\hat{y}_{12} = \hat{\beta}_1 X_2$

50

50

n = 100

$$MSE_{O.S} = \frac{1}{m} \sum_{1}^{m} (y_2 - \hat{y}_{12})^2$$

$$R^2_{O.S} = 1 - \frac{MSE_{OS}}{\frac{1}{m} \sum (y_2)^2}$$

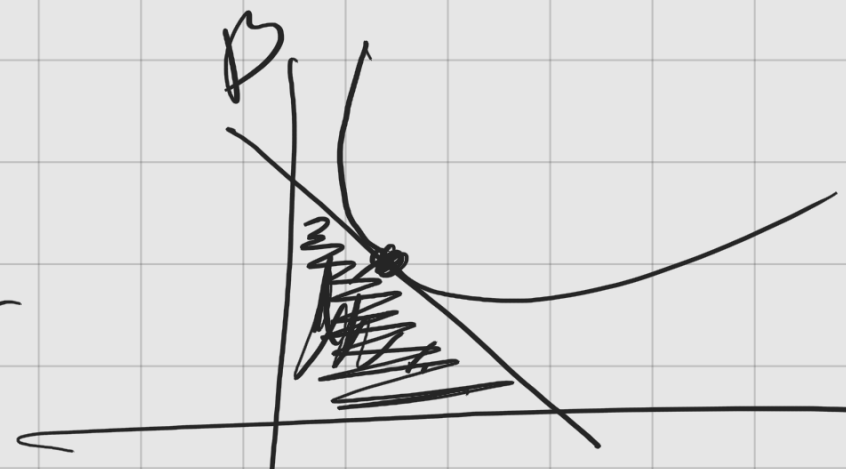
HDS / LASSO

* $\min_{b \in \mathbb{R}^p} \underbrace{\sum (y_i - b' x_i)^2}_{1} + \underbrace{\lambda \sum_{j=1}^{p} |b_j|}_{2}$

① MSE_p

② $\lambda \sum_{j=1}^{p} |b_j|$

* $\min G(K,L)$ S.t. $Y = f(K,L)$
* $\max U(A,B)$ ST $I = P_A A + P_B B$



$\max U(A,B) + \lambda (P_A A + P_B B - I)$

$$\min \sum_i (y_i - b' x_i)^2 \quad \text{s.t.} \quad \sum_{j=1}^{p} |b_j| \le s$$

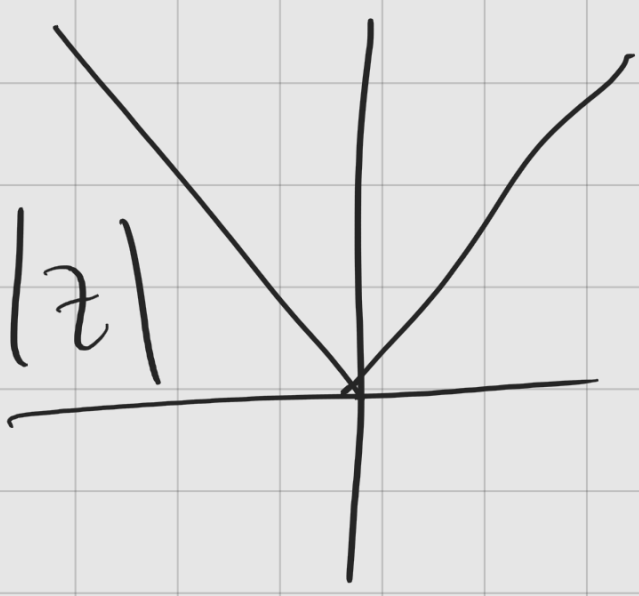
$$\min_b \sum_i (y_i - b' x_i)^2 + \lambda \left[\sum_{j=1}^{p} |b_j| - s \right]$$

$$\min_b \sum_i (y_i - b' x_i)^2 + \lambda \sum_{j=1}^{p} |b_j| \quad (1)$$

(2)

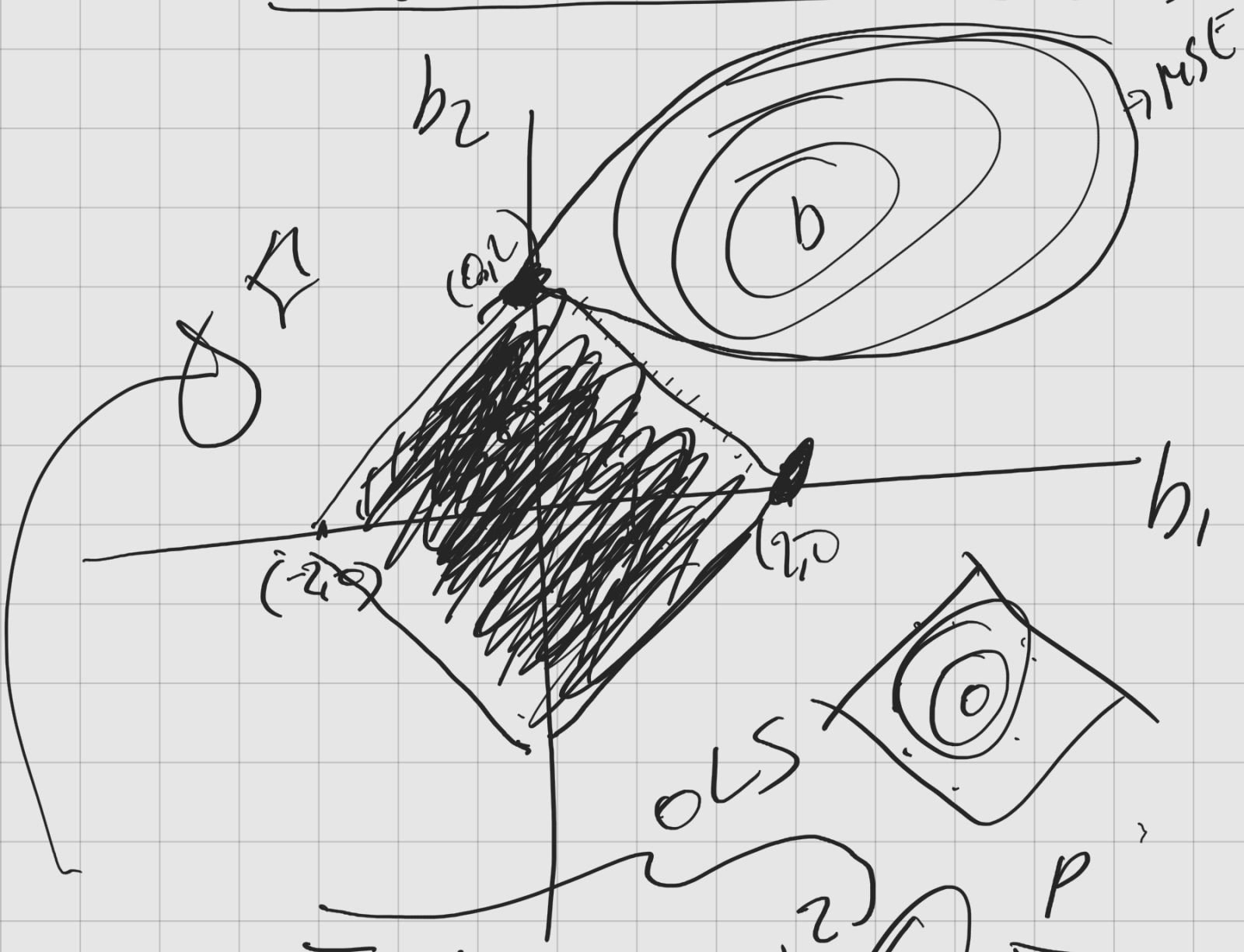
$x_1, x_2 \longrightarrow y$

$y = |x| + |z|$



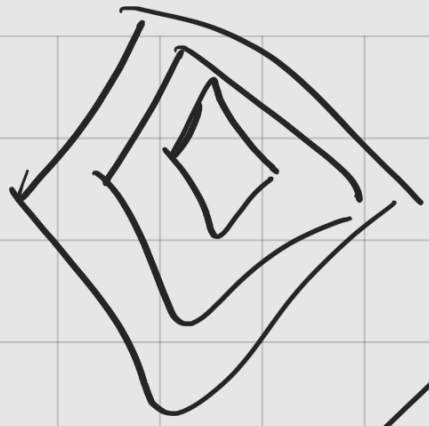
$\sum_{j=1}^{p} |b_j|$ $\boxed{p=2}$

$\hookrightarrow |b_1| + |b_2| \leq 2b$



$$\min_{b \in \mathbb{R}^p} \sum_i (y_i - b' x_i)^2 + \lambda \sum_{j=1}^{p} |b_j|$$

LASSO

$\lambda = 0 \cong$  OLS

$\lambda \uparrow \rightarrow b_1, b_2, b_3, b_4 \rightarrow$ OLS

b_1, b_2

Ex.

$\hat{y} = \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots \hat{\beta}_m x_m$

$\hat{y} = 1{,}5 x_1 + 2{,}3 x_2 + \dots 0{,}07 x_m$

$0{,}02 x_{m-1}$

$\lambda \uparrow \rightarrow$ ~~0,07~~ ~~0,02~~

$\lambda b \to 0 \longrightarrow$ 0,007 ✓
0,002 ✓

OLS

$n = 100$ ✓
$p = 100$ ✓ → OLS → overf.

$b = 0{,}003$

↳ LASSO ↙

↳ $p = 25$

* B, CH → 2013 *

$$\lambda = 2c\hat{\sigma}\sqrt{2n\log(2p/\delta)}$$

$$\hat{\sigma} = \sqrt{EE^2}$$

$c = 1{,}1$

$\gamma = 0{,}05 \to$ C.I

K-Fold CROSS VALIDATION

K = # Samples

1	2	3	4	5
6	7	8	9	10

$\lambda = \{ 0,01 ; 0,02 ; 0,0^2 \ldots 0,09 \}$

STEP 1

Fold 1 → TEST → MSE_1

2 . . . . . 10 → Train → $\hat{\beta}_2$

STEP 2:

Fold 2 → TEST → MSE_2

1 3 . . . . . 10 → TRAIN $\hat{\beta}_2$

.
.
.
.
.

STEP 10 $\longrightarrow$ $\sqrt{MSE_{10}}$

Ridge

$$\sum (y_i - \beta' x)^2 + \lambda_2 |b_i|^2$$

LASSO

$$\sum (y_i - \beta' x)^2 + \lambda_2 |b_i|$$

$$\sum (y_i - \beta' x)^2 + \alpha \lambda_1 |b_i|^2 +$$

$$(1-\alpha) \lambda_2 |b_i|$$

$$|b_1| + |b_2| + \dots |b_n| \leq S$$

$$|b_1|^2 + |b_2|^2 \dots |b_m|^2 \leq S$$