

# Examples of opt. in Stat/ML

Linear regression (penalized)

SVM

Logistic regression ..

Maximum likelihood

Mass transport

Gradient boosting

Model selection w/ AIC, BIC, RIC...

## Examples of contrary

Bootstrap/cross-validation / Monte Carlo

Bayesian inference ...

Forward stepwise regression ←

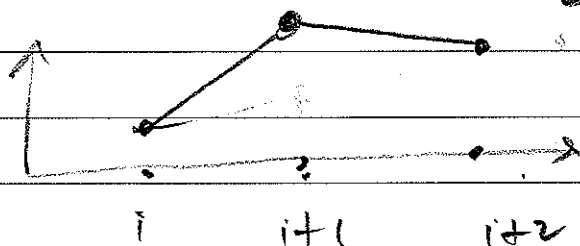
Clustering (Hierarchical) ←

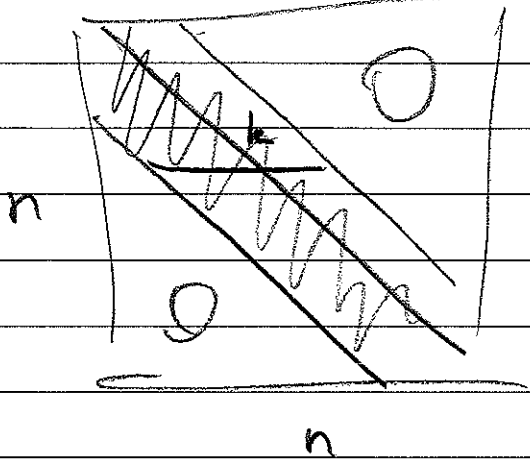
$$\frac{1}{2} \sum (y_i - \beta_i)^2 + \lambda \sum |\beta_i - 2\beta_{i+1} + \beta_{i+2}|$$

$$\beta_i - 2\beta_{i+1} + \beta_{i+2} = 0$$

$$\lambda \sum \mathcal{L}_i(\beta_i)$$

$$\Leftrightarrow \beta_{i+1} = \frac{\beta_i + \beta_{i+2}}{2}$$



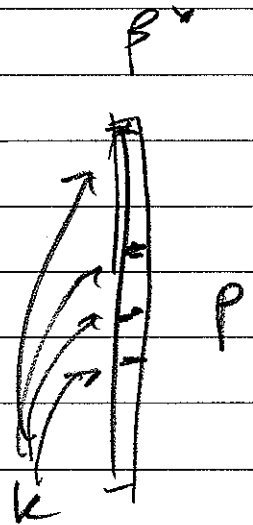
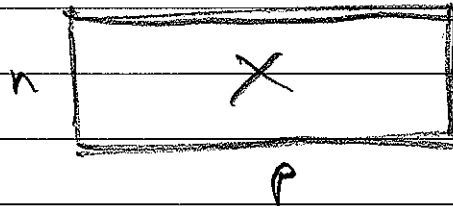


$$x = b$$

$$O(n^3)$$

$$O(nk^2)$$

$X$   $n \times p$   
 $n \times p$



$$y = X\beta^*$$

↑  
Sparse

$$r^{(0)} = y$$

$$\min \|\beta\|_1 \quad \text{s.t.} \quad \|y - X\beta\|_2 \leq \epsilon$$

$$h_j(x) = a_j^T x + b \quad \begin{array}{l} \text{affine} \\ \text{(linear)} \end{array}$$


---

Claim.  $x$  feasible

$$f(x) \leq f(y)$$

$$\|x - y\|_2 \leq \rho, \quad \rho > 0$$

$\Rightarrow f(x) \leq f(y)$  for all feas  $y$

i.e.  $x$  is a solution

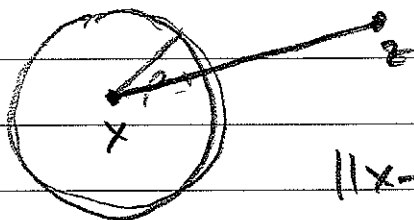
Proof. suppose  $\exists$  feasible  $z$  with  $f(z) < f(x)$

$y = (1-t)x + tz$  feasible for all  $t \in (0,1)$  ✓

$$\begin{aligned} g_i(y) &= g_i((1-t)x + tz) \\ &\leq (1-t)g_i(x) + tg_i(z) \\ &\leq 0. \end{aligned}$$

$$\begin{aligned} h_j(y) &= a_j^T((1-t)x + tz) + b_j \\ &= 0 \end{aligned}$$

$$\begin{aligned} f(y) &= f((1-t)x + tz) \\ &\leq (1-t)f(x) + tf(z) \\ &< f(x) \end{aligned}$$



$$\begin{aligned} \|x - y\|_2 &= \|x - tz - (1-t)x\|_2 \\ &= t\|x - z\|_2, \quad t = \rho / \|x - z\|_2 \end{aligned}$$