Examples

Regression (least squares) } + regularization
Classification (SVMs) } l_1 or l_2
Preference ordering (lasso or ridge)
Linear systems
$b = Ax$ or $y = X\beta$
↑ ↑
target model features

Resource allocation
Clustering (k-means)

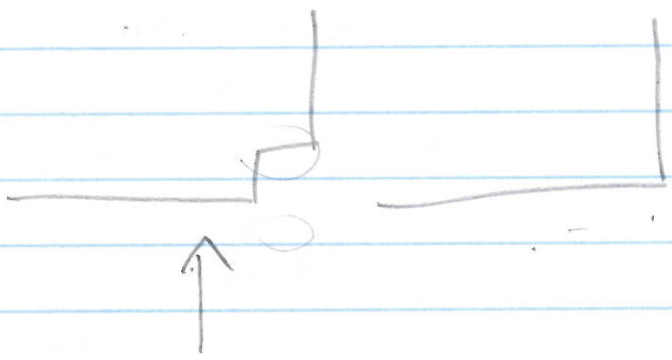
PCA

Non-examples

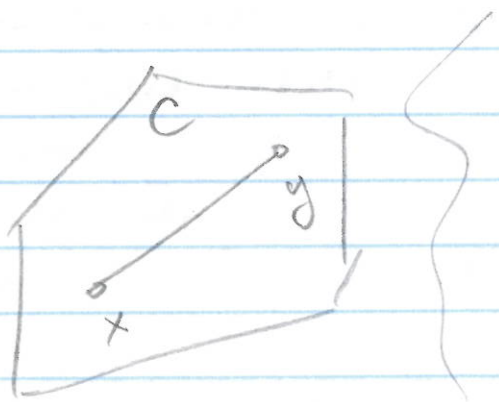
Hierarchical (clustering)
Density-based (clustering)
Tree-based models. ½
Boosting ½
Monte-Carlo
Cross-validation & Bootstrap

$$\sqrt{|\beta(i_x, i_y) - \beta(i_x+1, i_y)|^2 + |\beta(i_x, i_y) - \beta(i_x, i_y+1)|^2}$$

$$E = \{\{(i_x, i_y), (i_x', i_y')\} : |i_x - i_x'| = 1 \text{ or } |i_y - i_y'| = 1\}$$



$1 \{\beta_i \neq \beta_{i+1}\}$



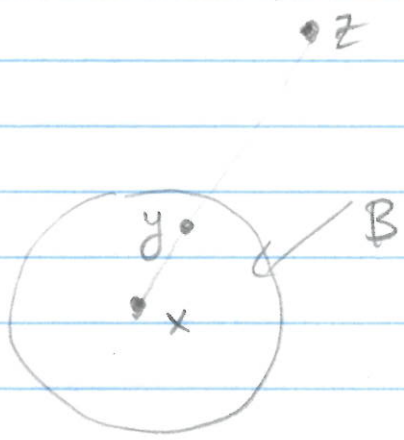
$\text{dom}(f) = \{x :$
$f(x)$ is defined
and finite $\}$

$h_j(x) = a_j^T x + b_j$ (affine)
(linear, $b_j = 0$)

$h_j(x) = 0 \iff h_j(x) \leq 0$
$-h_j(x) \leq 0$

Suppose not. assume there is some feasible z such that

$$f(z) < f(x)$$



Then $\|z - x\|_2 > \rho$

$y = tz + (1-t)x \quad t \in [0,1]$

- $y \in D$ ✓
- y satisfy constraints ✓

$$h_j(y) = a_j^T(tz + (1-t)x) + b_j$$
$$= 0$$

$$g_i(y) \leq t g_i(z) + (1-t) g_i(x)$$
$$\leq 0$$

take t small enough so that $y \in B$

$$t > 0$$

$$f(y) \leq t f(z) + (1-t) f(x)$$
$$< f(x)$$

× contradiction