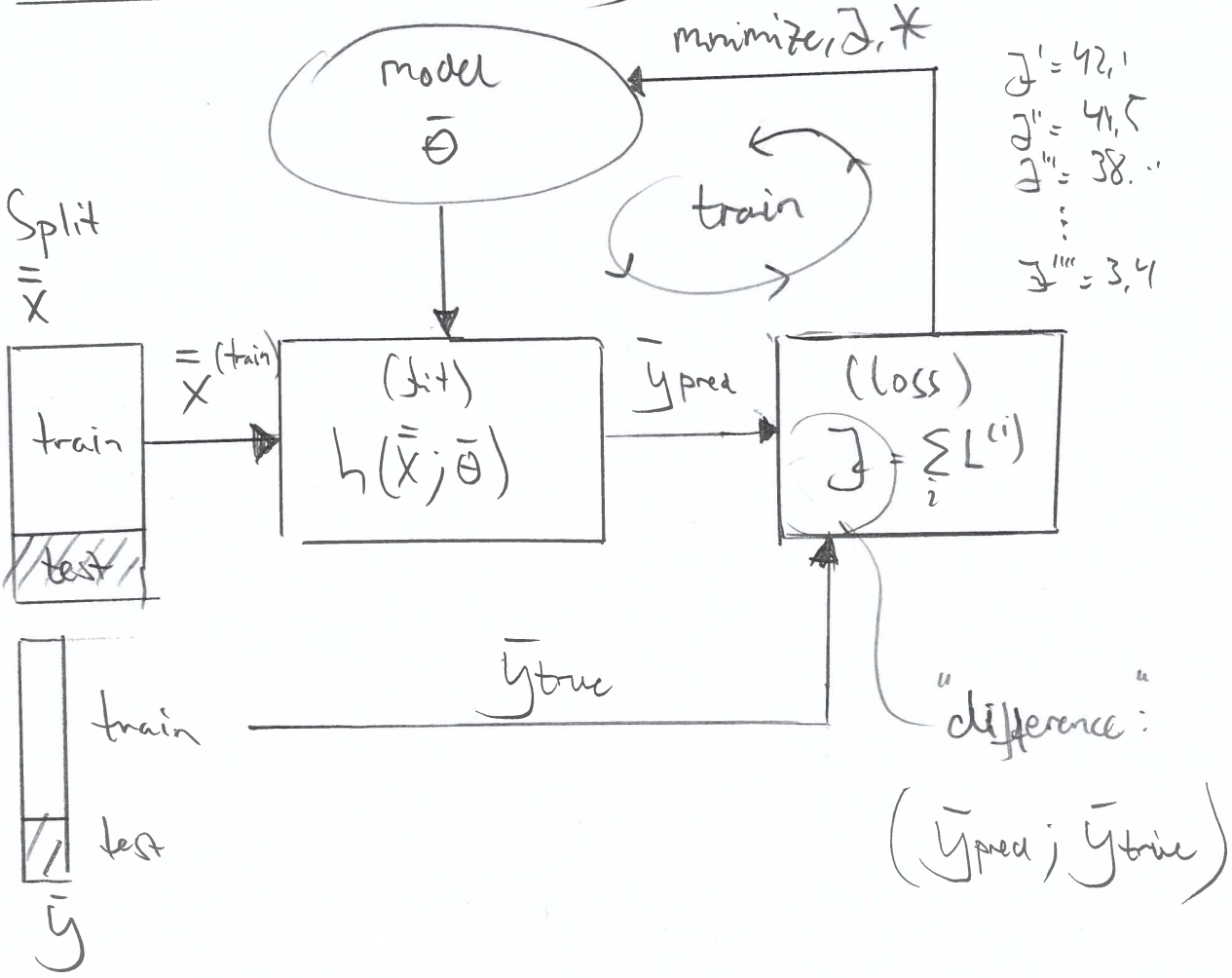
ITMAL: L02

Supervised Learning



MMLS mxn matrix

$$\bar{\bar{X}} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix}$$

ML: nxd matrix

$$\bar{\bar{X}} = \begin{bmatrix} \bar{x}^{(1)T} \\ \bar{x}^{(2)T} \\ \vdots \\ \bar{x}^{(n)T} \end{bmatrix} = \begin{bmatrix} x_1^{(1)} & x_2^{(1)} & \cdots & x_d^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \cdots & x_d^{(2)} \\ \vdots & & \ddots & \\ x_1^{(n)} & x_2^{(n)} & \cdots & x_d^{(n)} \end{bmatrix}$$

Cost / Distance / Norms

J: loss fun.
$L^{(i)}$: individual loss fun $\Big\}$ $J = \sum_i L^{(i)}$

$L^{(i)}(\bar{y}^{(i)}_{pred}, \bar{y}^{(i)}_{true}) \stackrel{?}{=} \dots$

Euclidian Norm

$\mathcal{L}_2$: $\|\bar{x}\|_2 = \sum_j |x_j|^2$

Ehs

$$\bar{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \quad \|x\|_2 \stackrel{?}{=} \sqrt{|1|^2 + |3|^2} = \sqrt{10}$$

Linear Algebra

$$\mathcal{L}_2^2: \|\bar{x}\|_2^2 = \bar{x}^T \bar{x}$$

Ehs

$$\|x\|_2^2 \stackrel{?}{=} \begin{bmatrix} 1 \\ 3 \end{bmatrix}^T \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = 1^2 + 3^2 = 10$$

$(m \times n = 1 \times 1)$

Distance $\bar{x}$ to $\bar{y}$

$$d(\bar{x};\bar{y}) = \|\bar{x}-\bar{y}\|_2 = \|\bar{z}\|_2 \qquad (\bar{z}=\bar{x}-\bar{y})$$

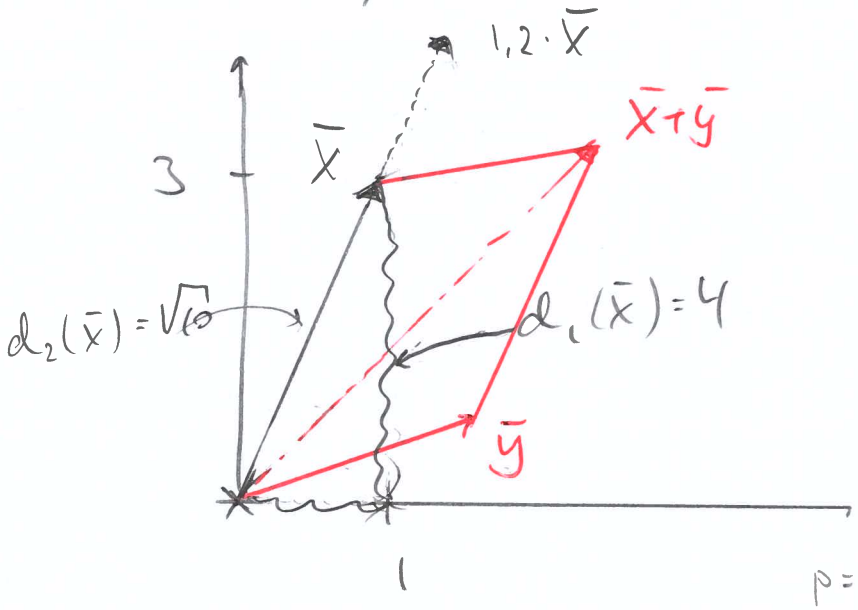
General norm

$$\mathcal{L}_p: \quad \|x\|_p = \left(\sum_j |x_j|\right)^{1/p}$$

Fels. City norm (Manhattan-norm)

$$\mathcal{L}_1: \quad \|\bar{x}\|_1 = \left\|\begin{bmatrix}1\\3\end{bmatrix}\right\|_1 \overset{?}{=} |1|+|3| = 4$$

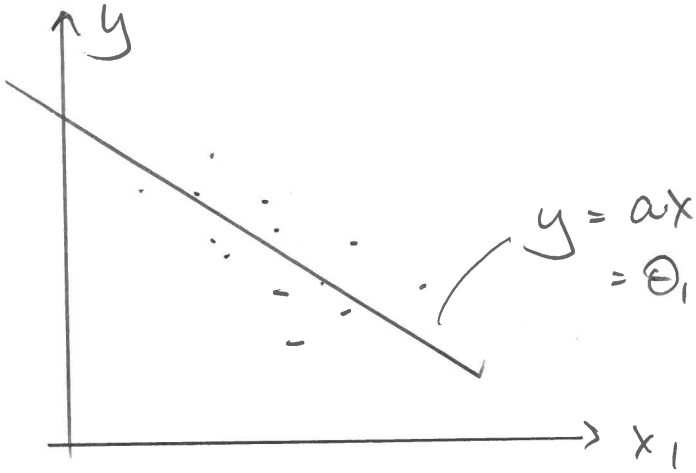
Distances / Norms



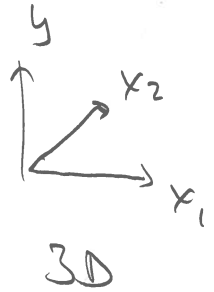
$p=2, \bar{x}=\begin{bmatrix}1\\3\end{bmatrix}$

$$d_p(\bar{0}) = 0$$
$$d_p(\alpha\bar{x}) = |\alpha|\, d_p(\bar{x}) = |1{,}2|\cdot\sqrt{10}$$
$$d_p(\bar{x}+\bar{y}) \le d_p(\bar{x}) + d_p(\bar{y})$$

Linear Regression



feature space, 2D



2D:

$$y_{pred} = \theta_1 x_1 + \theta_0 \cdot 1$$

General, N-D:

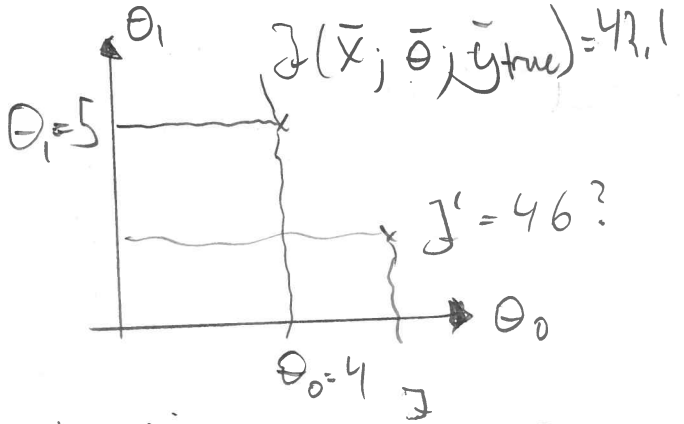
$$\boxed{h(\bar{x}; \bar{\theta}) = \bar{\theta}^T \cdot \bar{x}} = y_{pred}$$

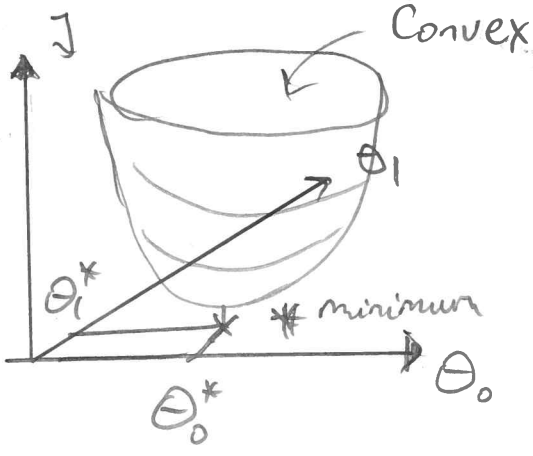
1D 2D

$$\theta = \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix} = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix}$$

$$x = \begin{bmatrix} 1 \\ x_1 \end{bmatrix} = \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix}$$

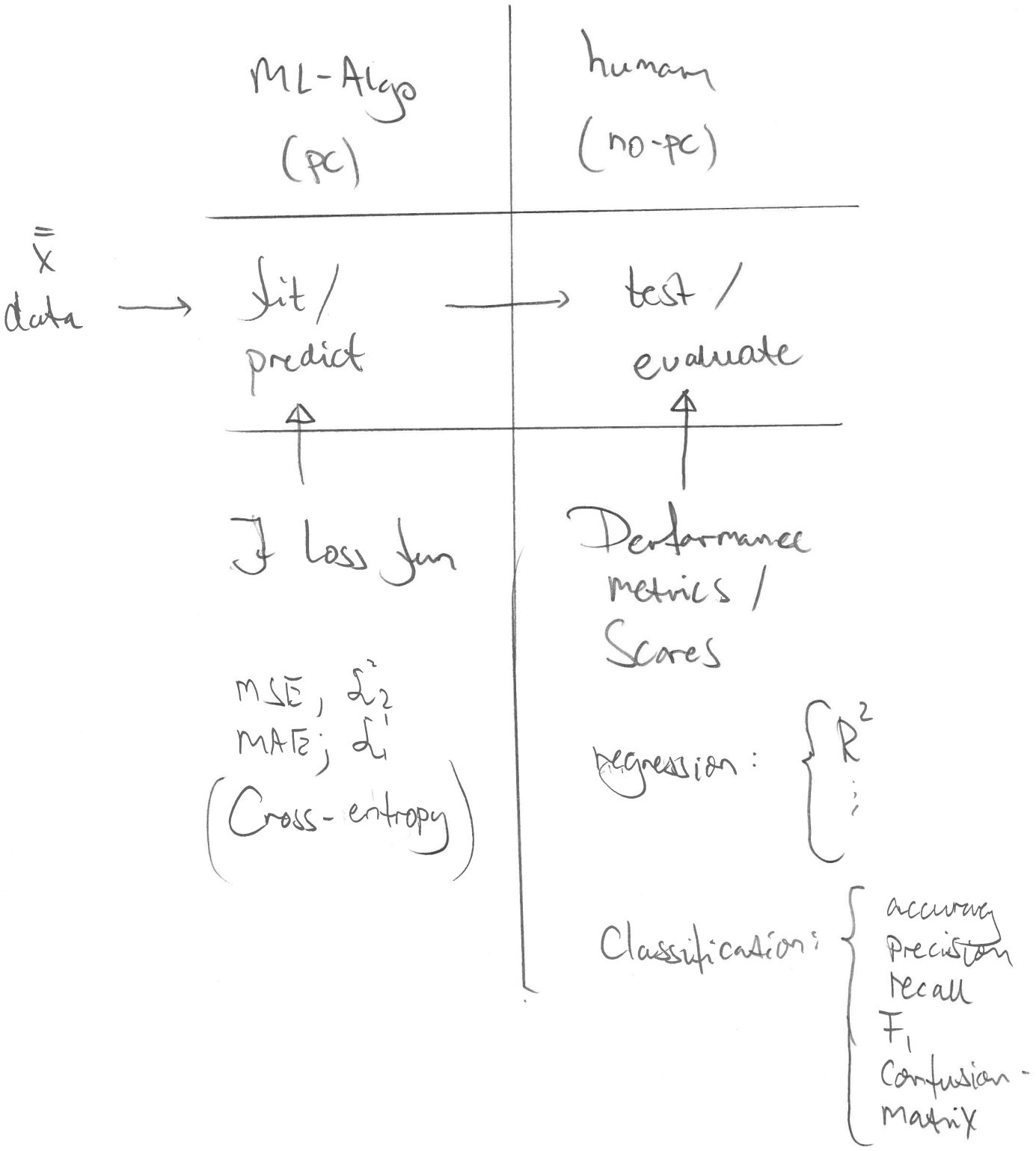
Parameter-Space





N-dim:

Performance Metrics



Binary Classification

Iris Setosa
Iris versicolor
} 2 classes ("binary")

⇒ Binary Classifikator? { Setosa → P
no-setosa → N

Confusion matrix?

		y_{true} P	y_{true} N
y_{pred}	P	TP	FP (I)
	N	FN (II)	TN