Regression

ML - Supervised Learning

Dan Calderone - Win22

Data

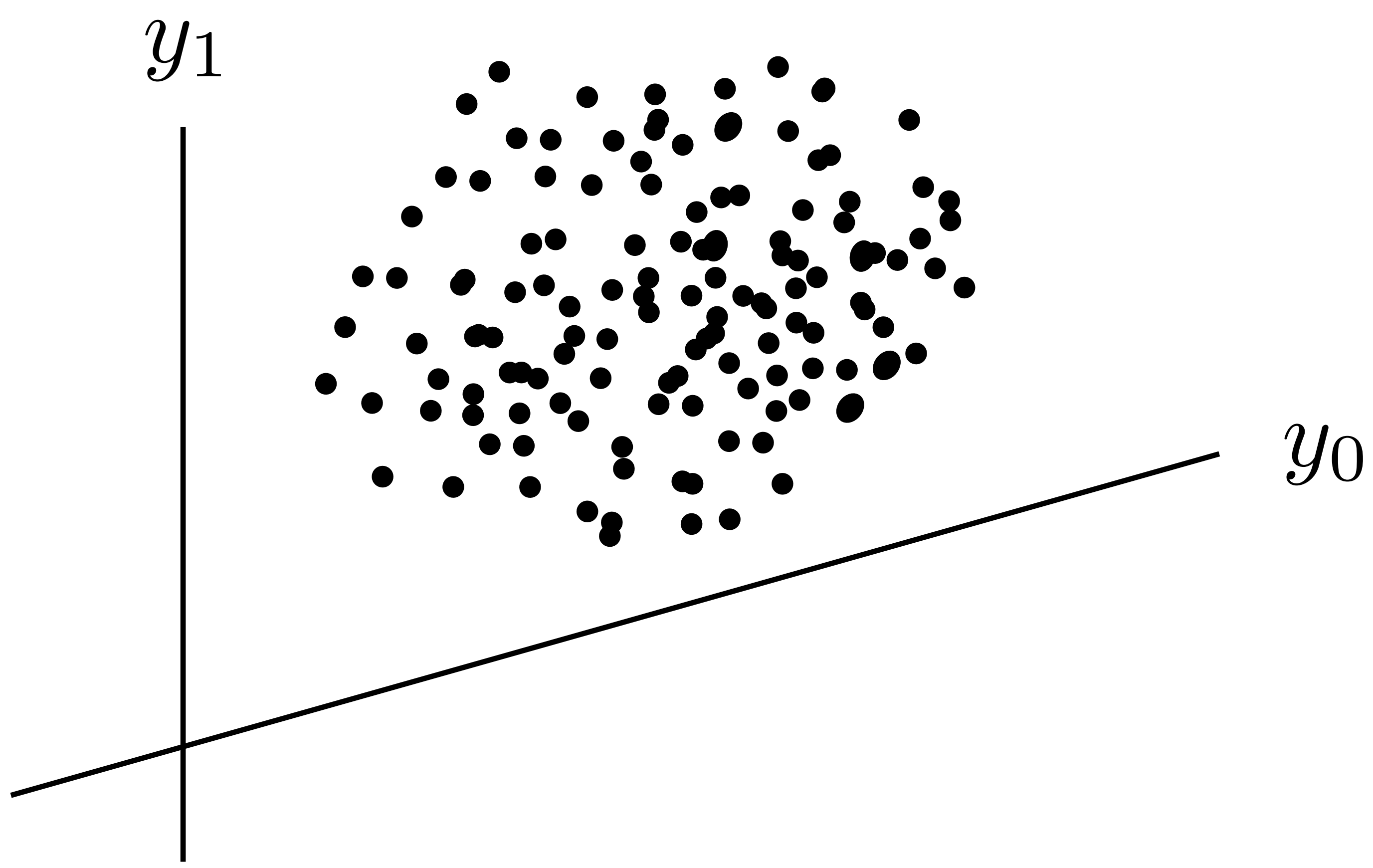
OUTPUTS
(Dependent Variables)

$$y_t = f(x_t)$$

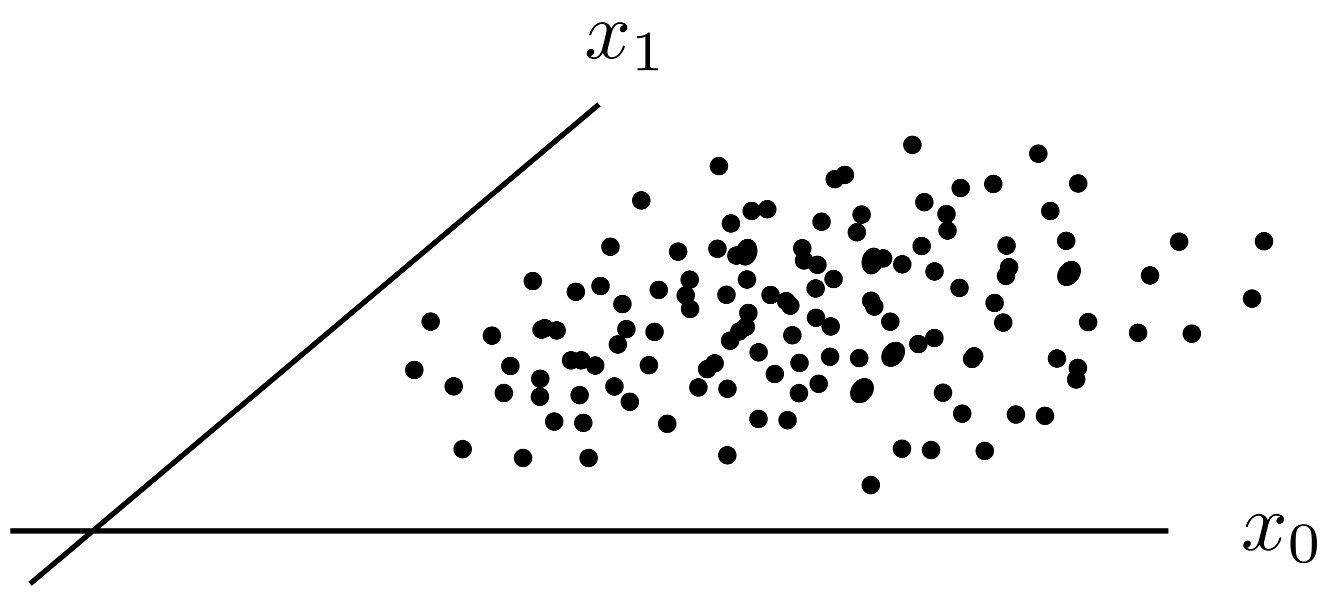
INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

continuous
variables



continuous
variables



Data

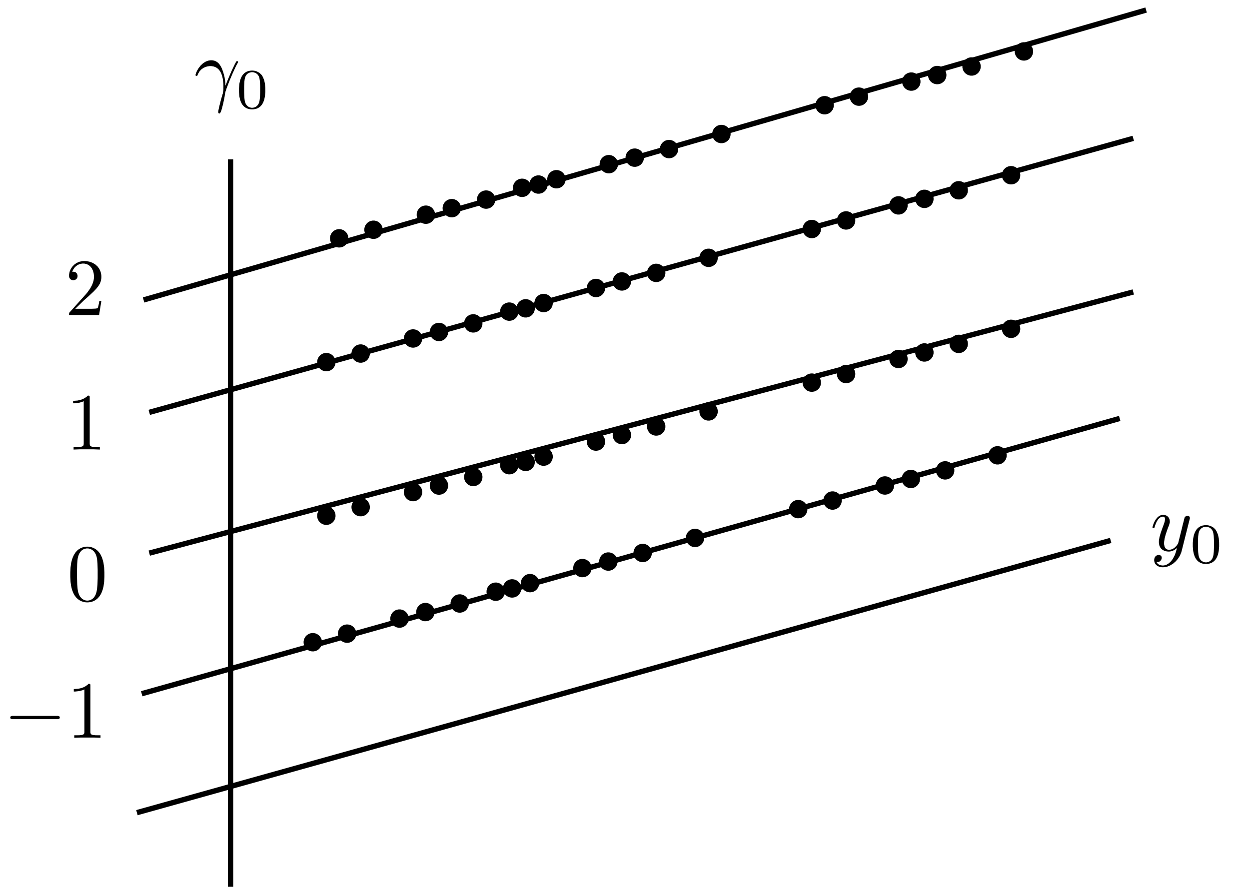
OUTPUTS (Dependent Variables)

$$y_t = f(x_t)$$

INPUTS (Independent Variables)

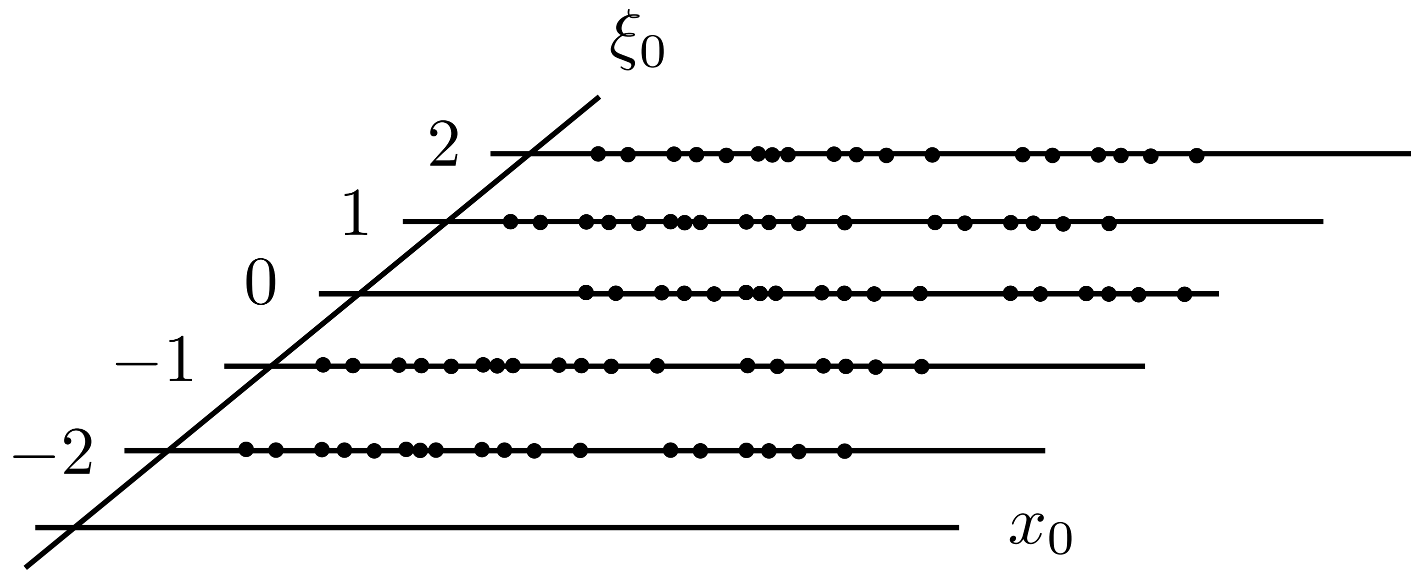
$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

…and/or discrete variables



Integer valued

…and/or discrete variables



Data

OUTPUTS
(Dependent Variables)

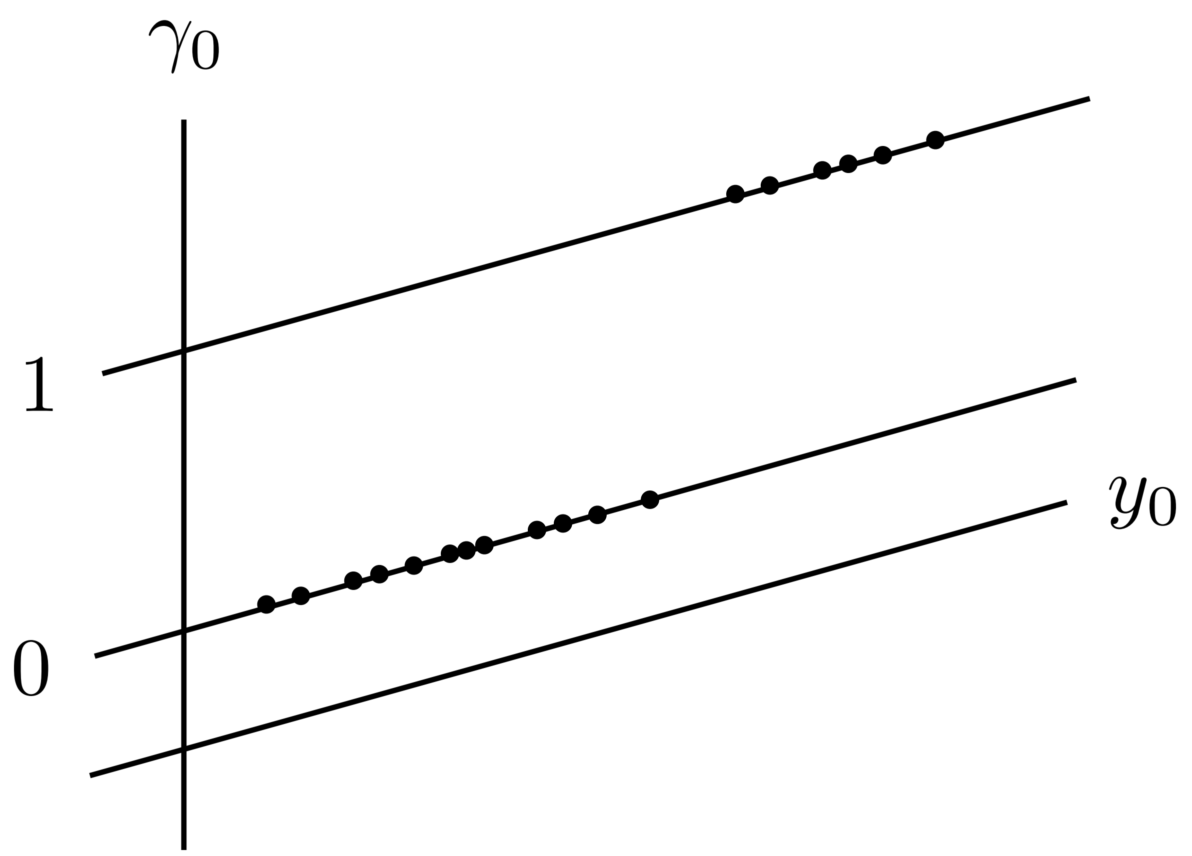
$$y_t = f(x_t)$$

INPUTS
(Independent Variables)

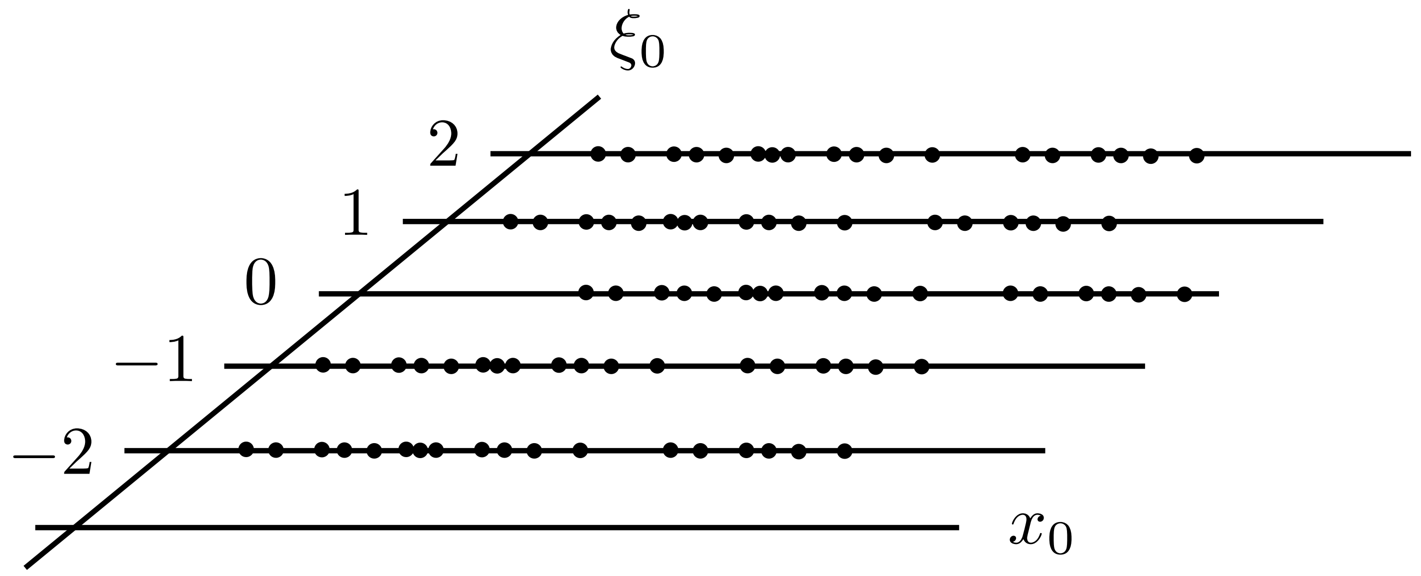
$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

…and/or discrete variables

…and/or discrete variables



boolean



Basis Functions

OUTPUTS (Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS (Independent Variables)

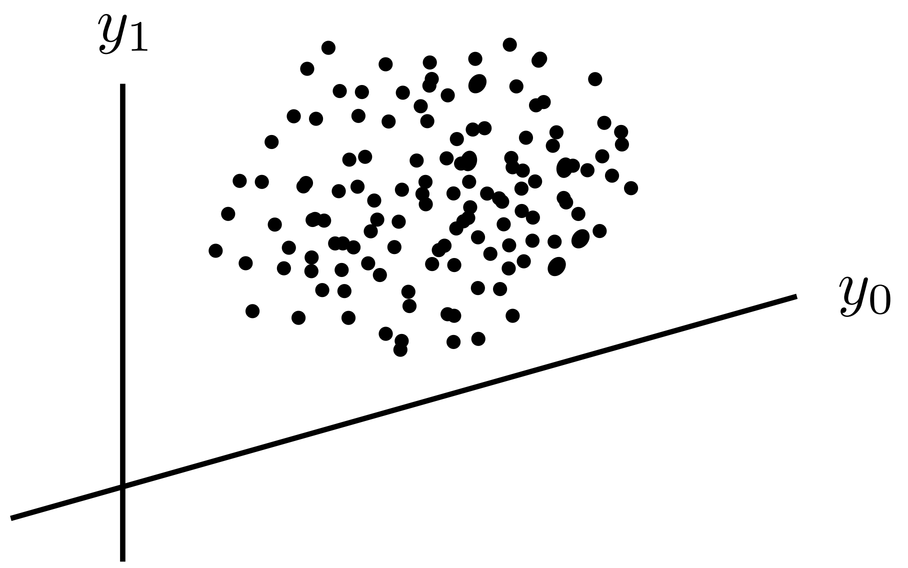
$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

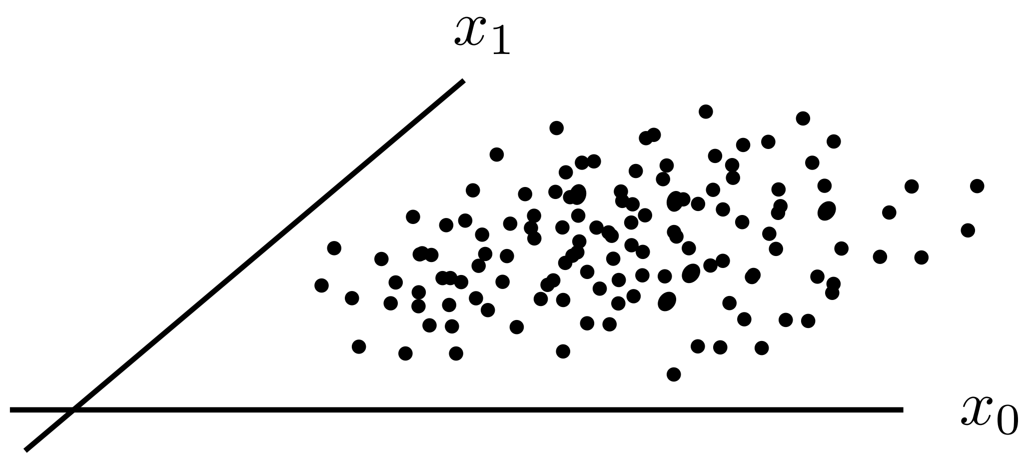
BASIS FUNCTIONS

polynomials…

exponentials…

Fourier basis





Basis Functions

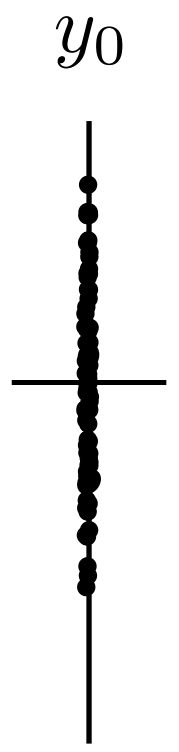
$$y_t = f(h_t(x_t, \xi_t))$$

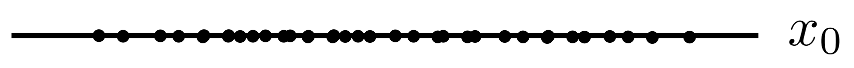
OUTPUTS (Dependent Variables) $\Leftarrow$ **BASIS FUNCTIONS** $\Leftarrow$ **INPUTS (Independent Variables)**

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

BASIS FUNCTIONS

$$h_t(x_t) = [1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3]$$





Basis Functions

$$y_t = f(h_t(x_t, \xi_t))$$

OUTPUTS
(Dependent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow$$

BASIS
FUNCTIONS

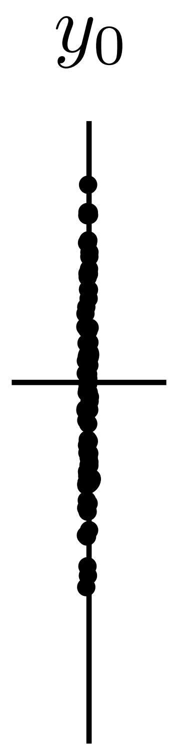
$$\begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Longleftarrow$$

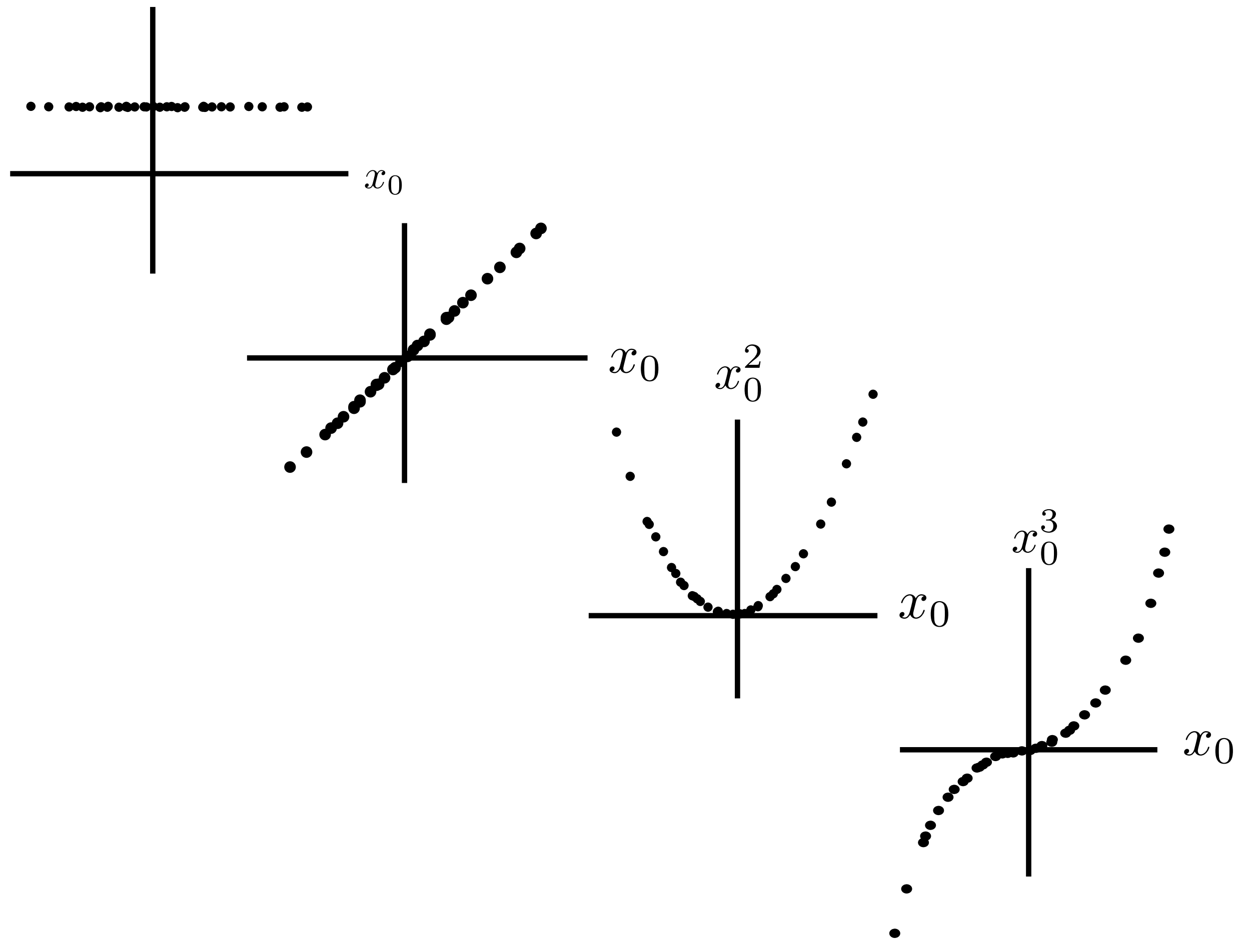
INPUTS
(Independent Variables)

$$\begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

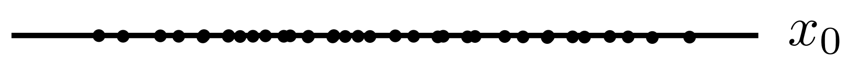
BASIS FUNCTIONS

$$h_t(x_t) = [1 \ x_{t0} \ x_{t0}^2 \ x_{t0}^3]$$





$\Longleftarrow$



Basis Functions

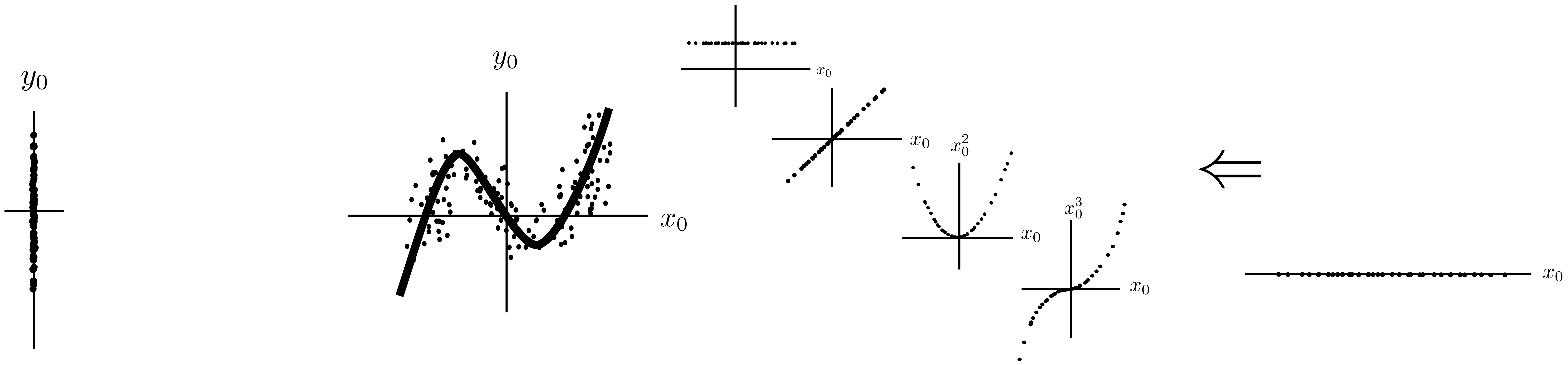
$$y_t = f(h_t(x_t, \xi_t))$$

OUTPUTS (Dependent Variables) | **BASIS FUNCTIONS** | **INPUTS (Independent Variables)**

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

BASIS FUNCTIONS

$$h_t(x_t) = [1 \ x_{t0} \ x_{t0}^2 \ x_{t0}^3]$$



Discrete Outputs

$$y_t = f(h_t(x_t, \xi_t))$$

OUTPUTS (Dependent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} \\ \gamma_{10} & \cdots & \gamma_{1m'} \\ \gamma_{20} & \cdots & \gamma_{2m'} \\ \gamma_{30} & \cdots & \gamma_{3m'} \\ \gamma_{40} & \cdots & \gamma_{4m'} \\ \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} \end{bmatrix} \Bigg| \begin{bmatrix} y_{00} & \cdots & y_{0m} \\ y_{10} & \cdots & y_{1m} \\ y_{20} & \cdots & y_{2m} \\ y_{30} & \cdots & y_{3m} \\ y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots \\ y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow$$

Discrete Output | Thresh-hold | Continuous Prediction

BASIS FUNCTIONS

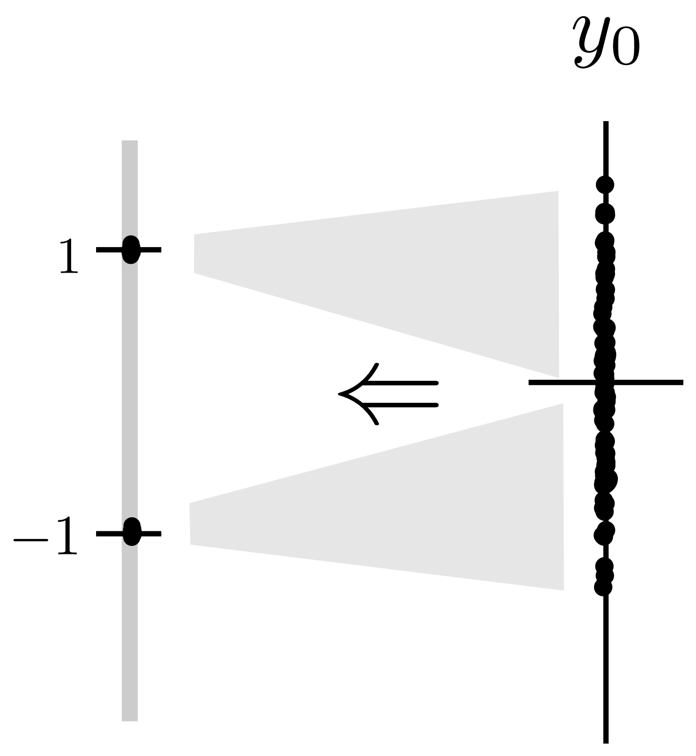
$$\begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Longleftarrow$$

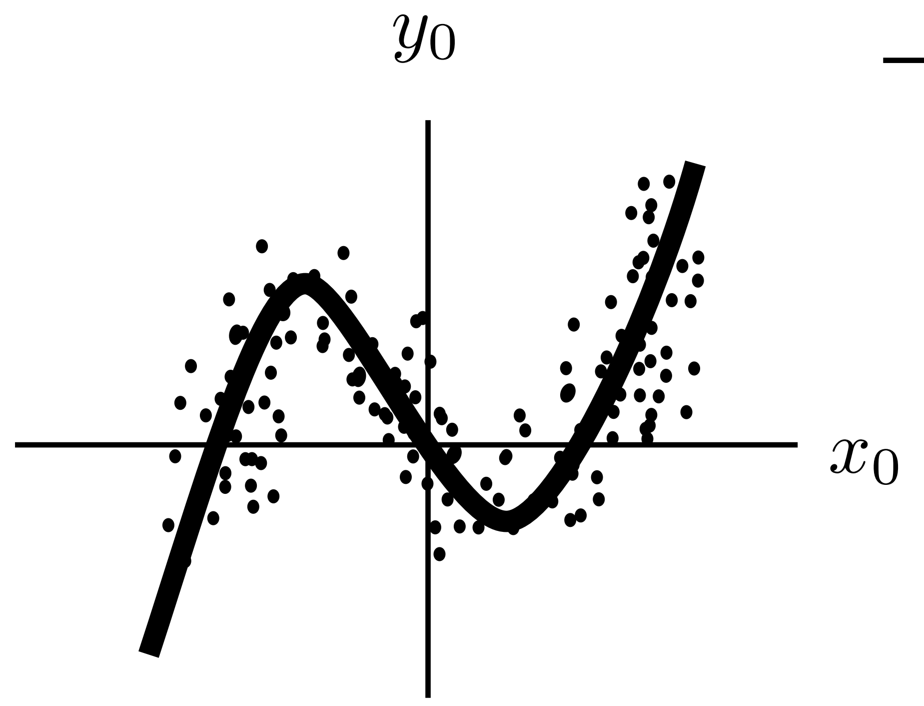
INPUTS (Independent Variables)

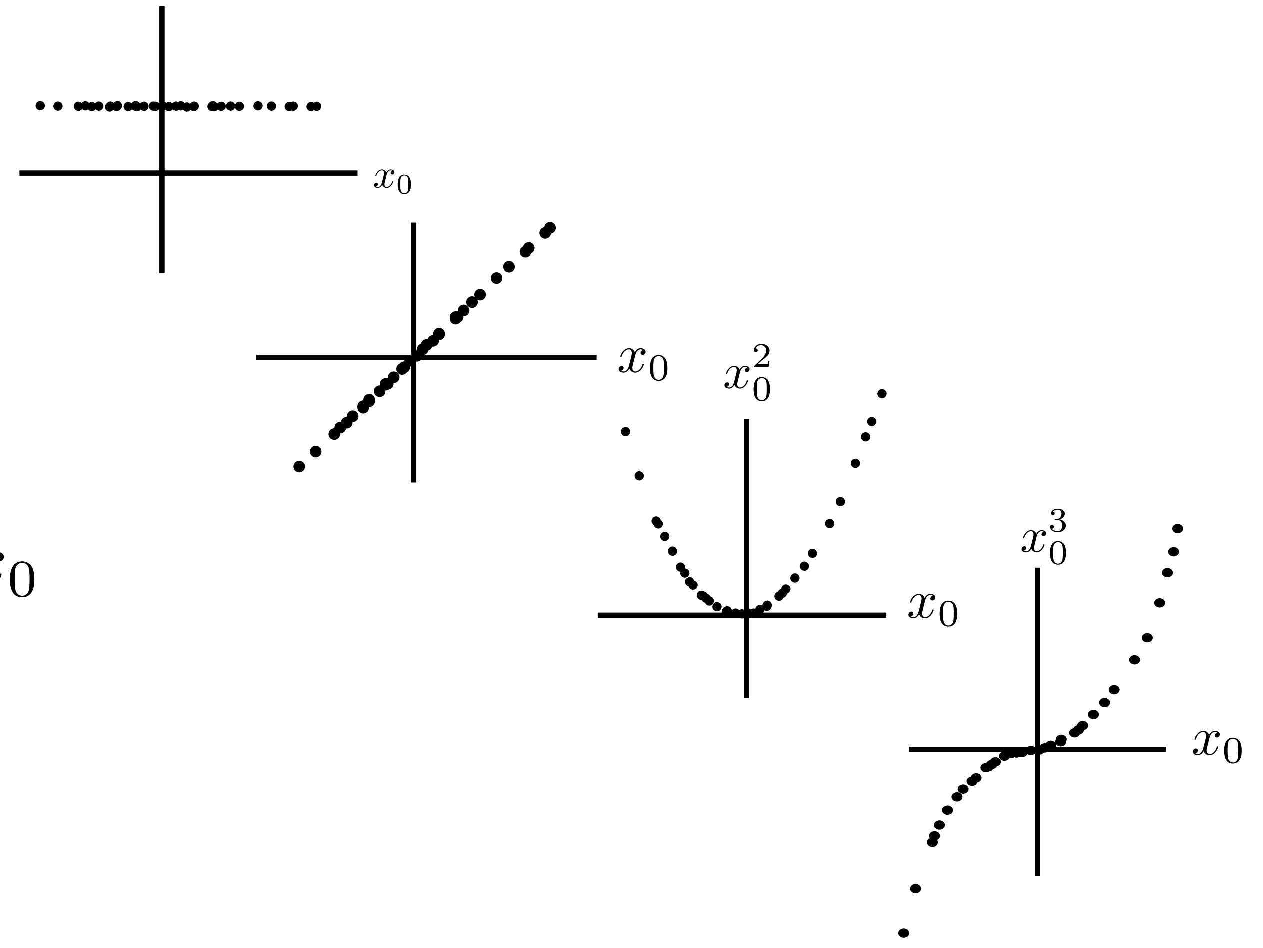
$$\begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

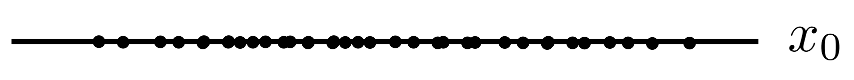
BASIS FUNCTIONS

$$h_t(x_t) = [1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3]$$









Loss Functions

OUTPUTS
(Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

BASIS FUNCTIONS

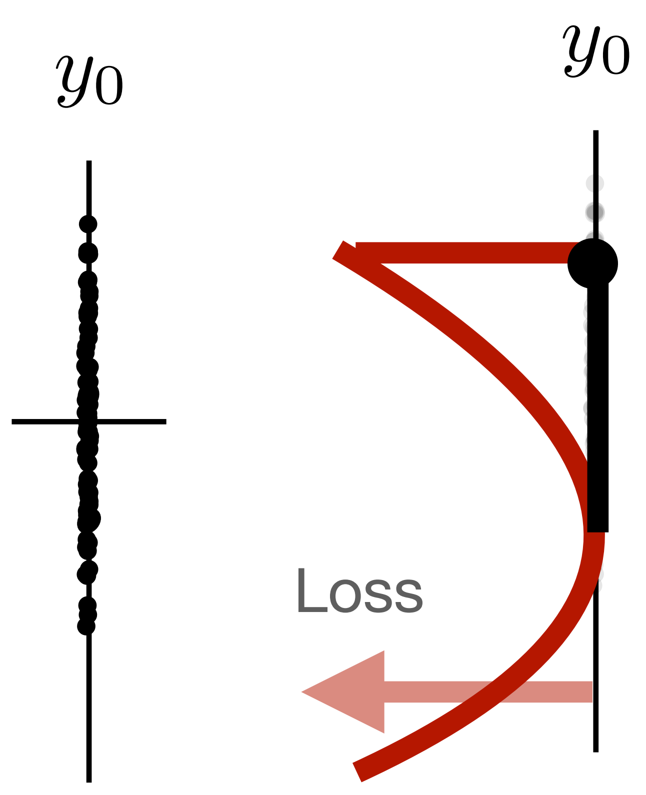
Mean squared error

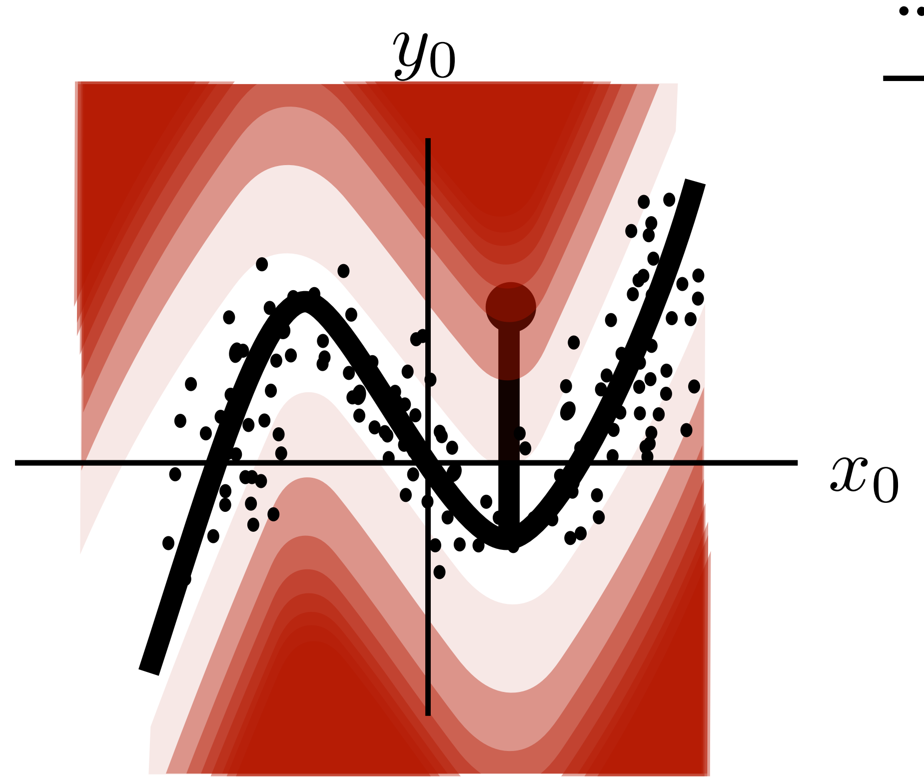
Quadratic

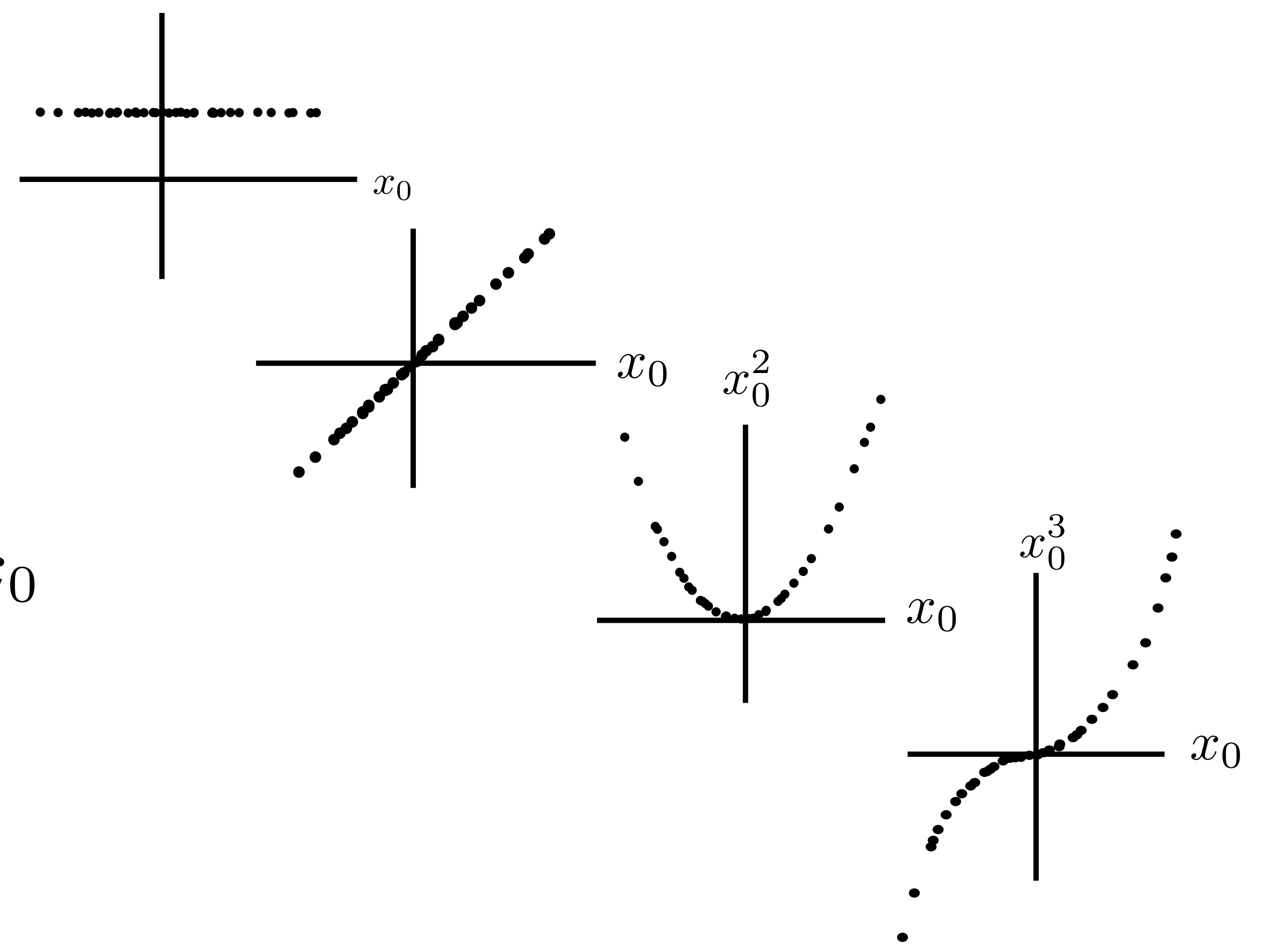
$$||\cdot||_2^2$$

$$\left|\left|y - f(h(x, \xi))\right|\right|_2^2$$

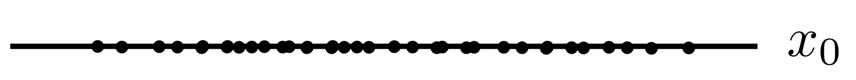
$$h_t(x_t) = [1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3]$$







$\Leftarrow$



Loss Functions

OUTPUTS (Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS (Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

BASIS FUNCTIONS

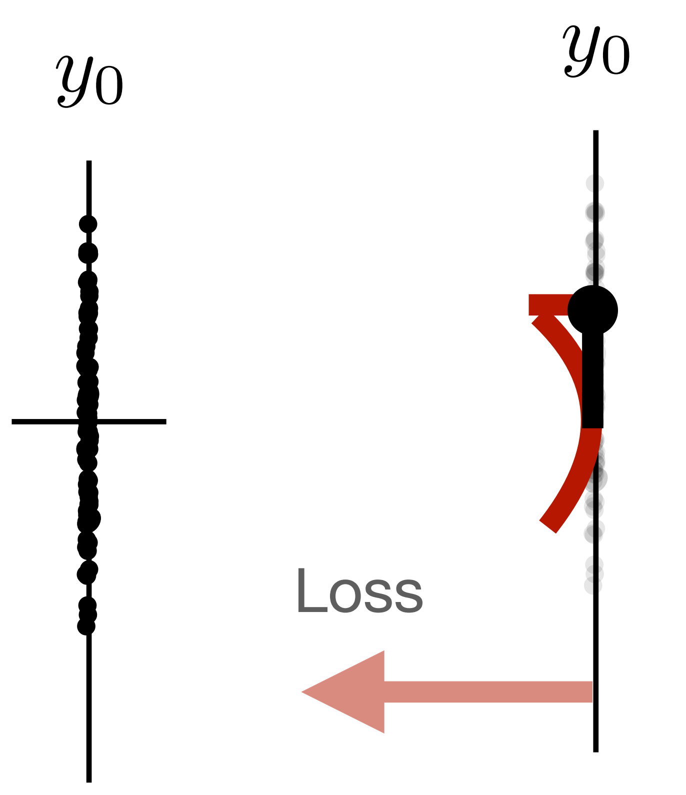
Mean squared error

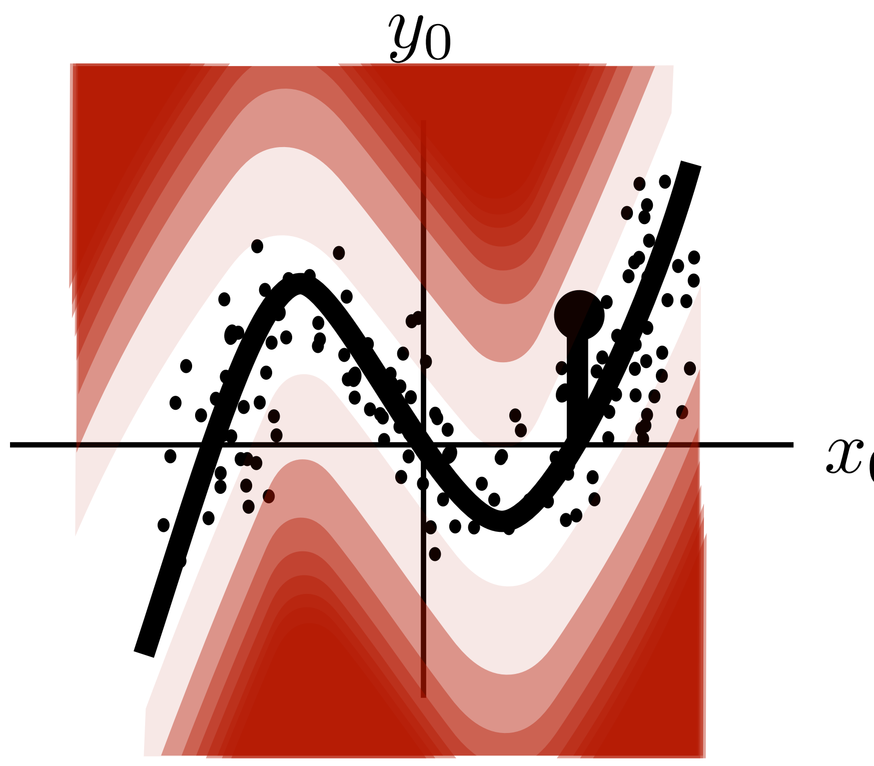
Quadratic

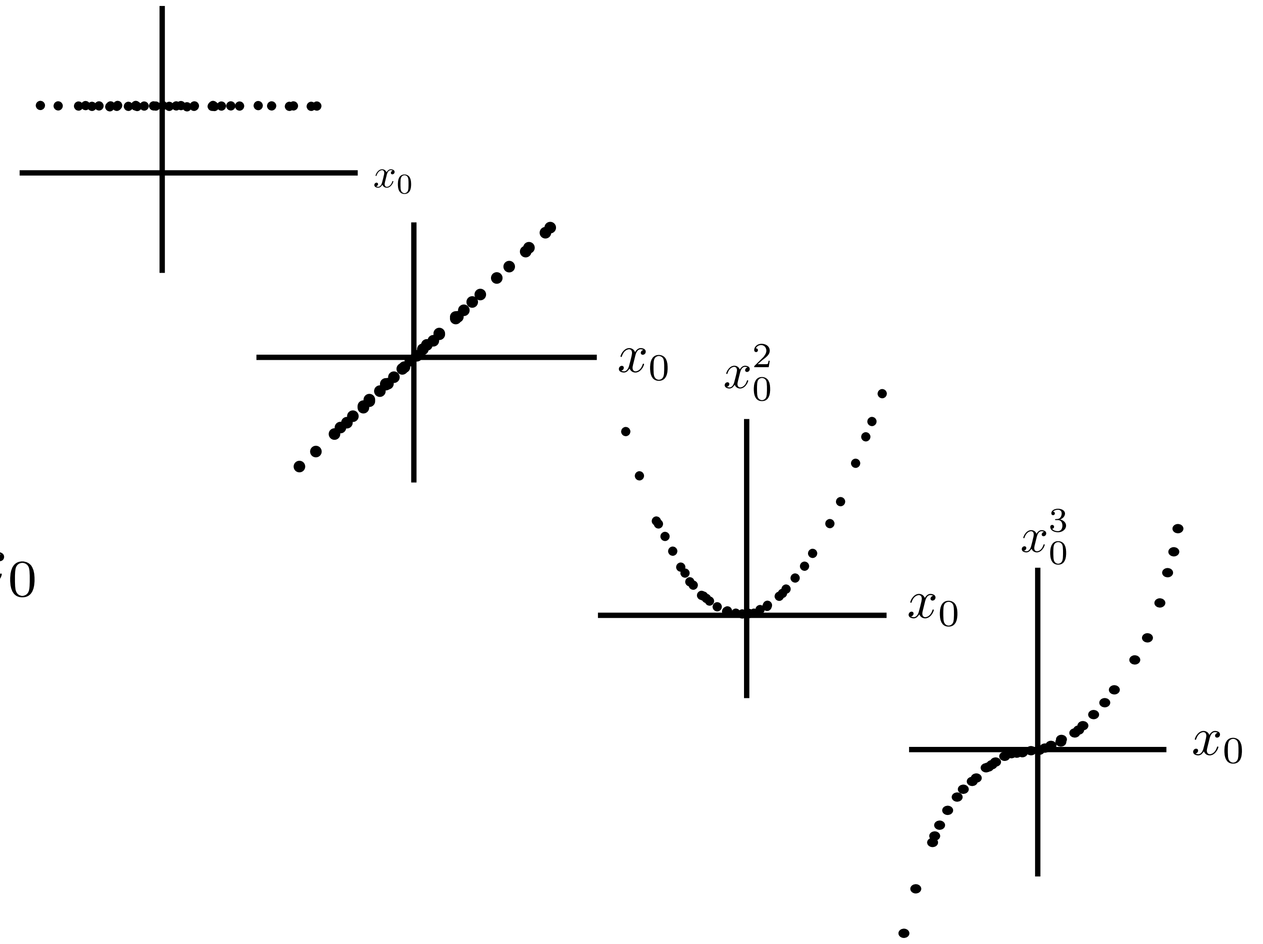
$$||\cdot||_2^2$$

$$\left|\left|y - f(h(x, \xi))\right|\right|_2^2$$

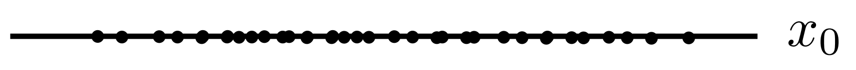
$$h_t(x_t) = [1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3]$$







$\Leftarrow$



Loss Functions

OUTPUTS (Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS (Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Longleftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

$$\left|\left|y - f(h(x, \xi))\right|\right|_2^2$$

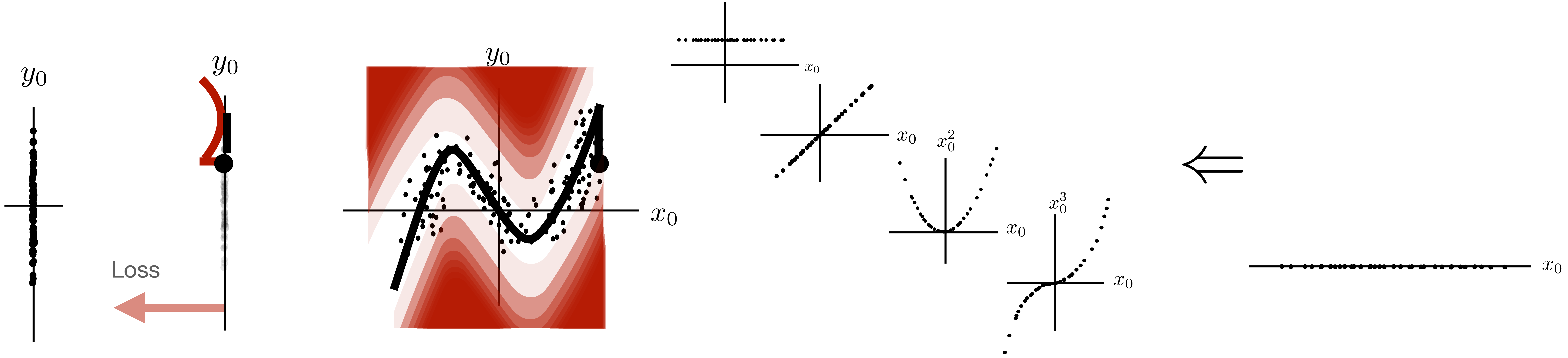
BASIS FUNCTIONS

$$h_t(x_t) = [1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3]$$

Mean squared error

Quadratic

$$||\cdot||_2^2$$



Loss Functions

OUTPUTS
(Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0,\xi_0) & \cdots & h_{0n}(x_0,\xi_0) \\ h_{10}(x_1,\xi_1) & \cdots & h_{1n}(x_1,\xi_1) \\ h_{20}(x_2,\xi_2) & \cdots & h_{2n}(x_2,\xi_2) \\ h_{30}(x_3,\xi_3) & \cdots & h_{3n}(x_3,\xi_3) \\ h_{40}(x_4,\xi_4) & \cdots & h_{4n}(x_4,\xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T,\xi_T) & \cdots & h_{Tn}(x_T,\xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

BASIS FUNCTIONS

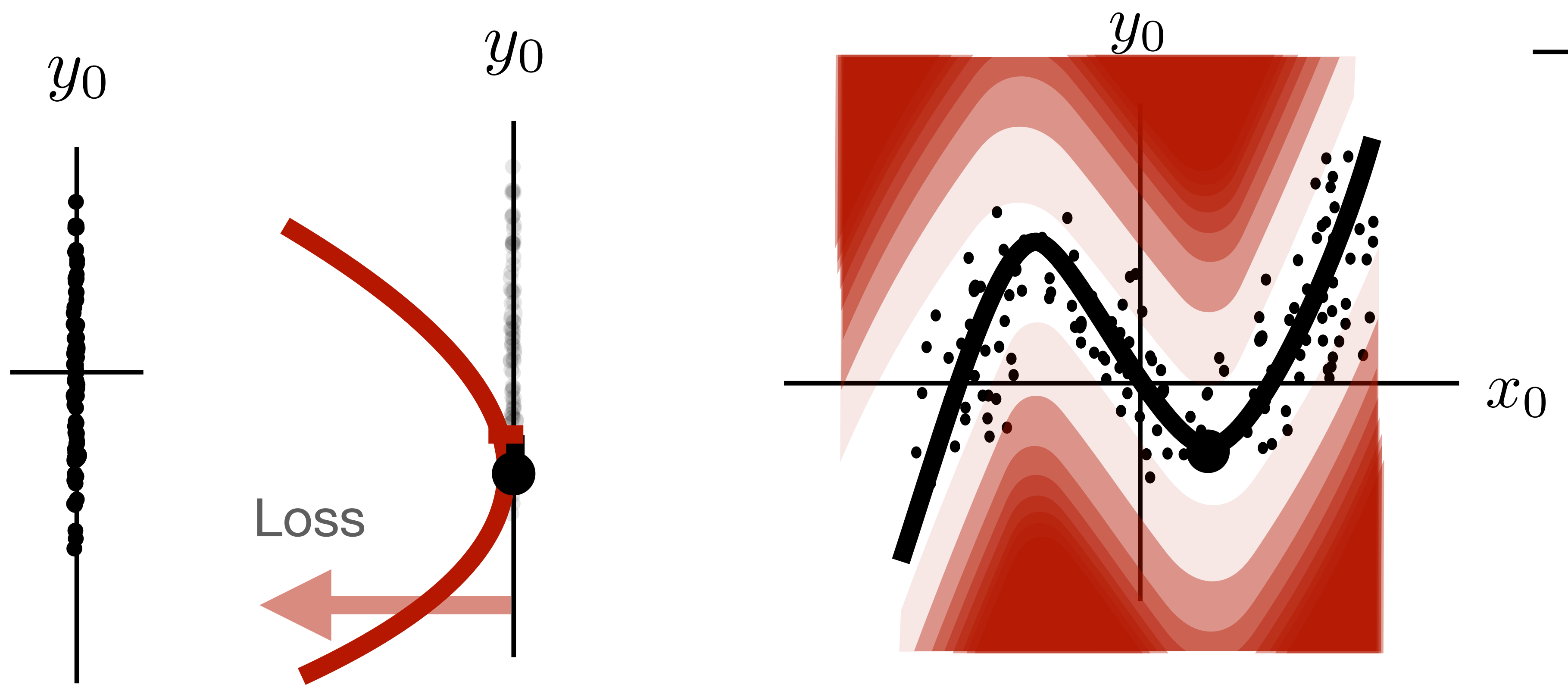
Mean squared error

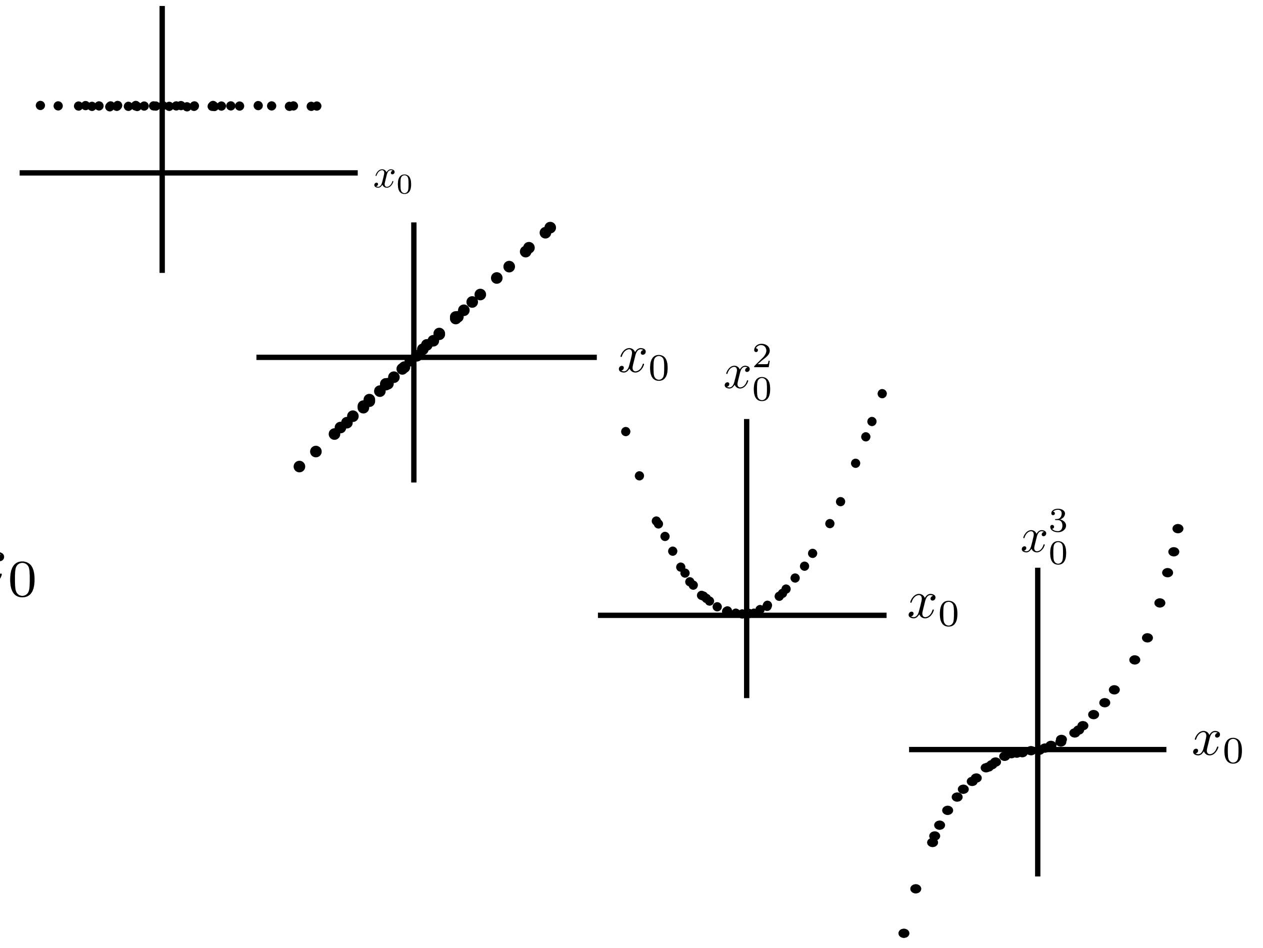
Quadratic

$||\cdot||_2^2$

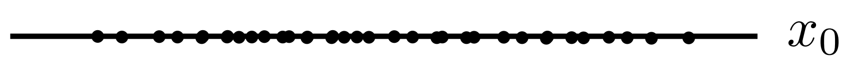
$$\Big|\Big|y - f(h(x,\xi))\Big|\Big|_2^2$$

$$h_t(x_t) = [1 \ x_{t0} \ x_{t0}^2 \ x_{t0}^3]$$





$\Leftarrow$



Loss Functions

OUTPUTS (Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS (Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

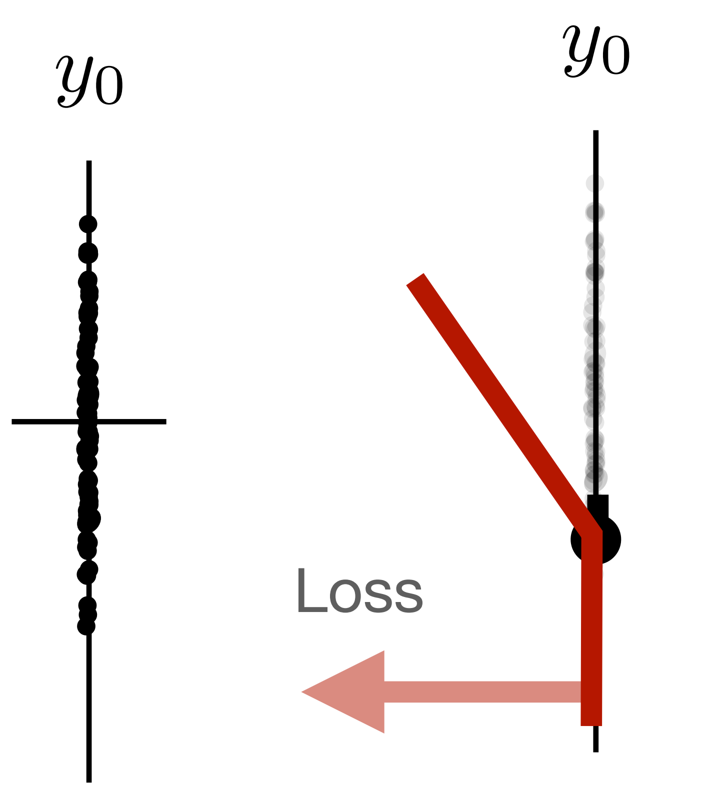
BASIS FUNCTIONS

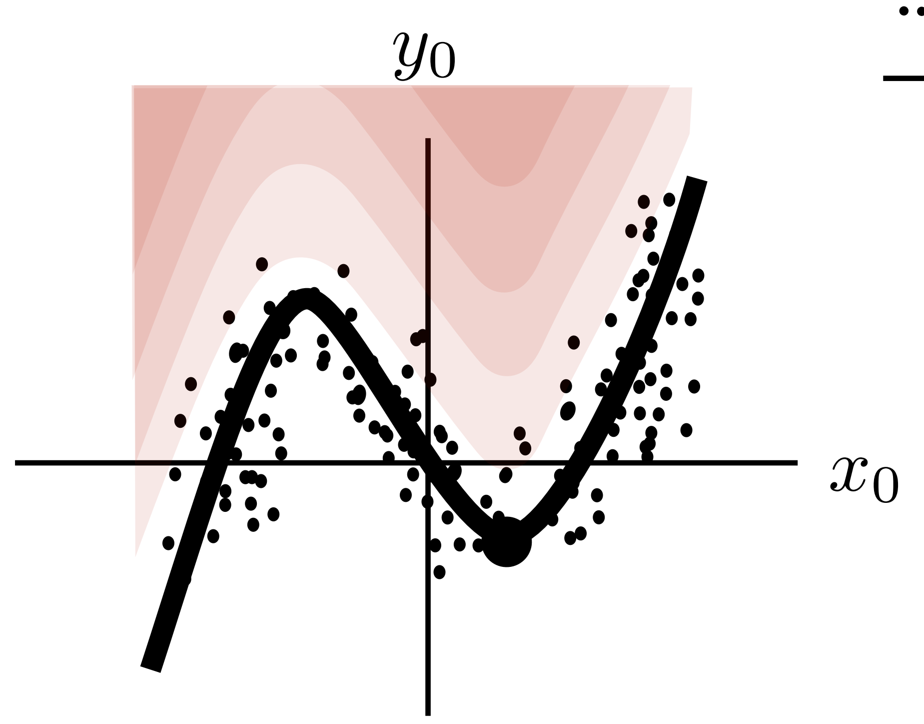
Hinge Loss

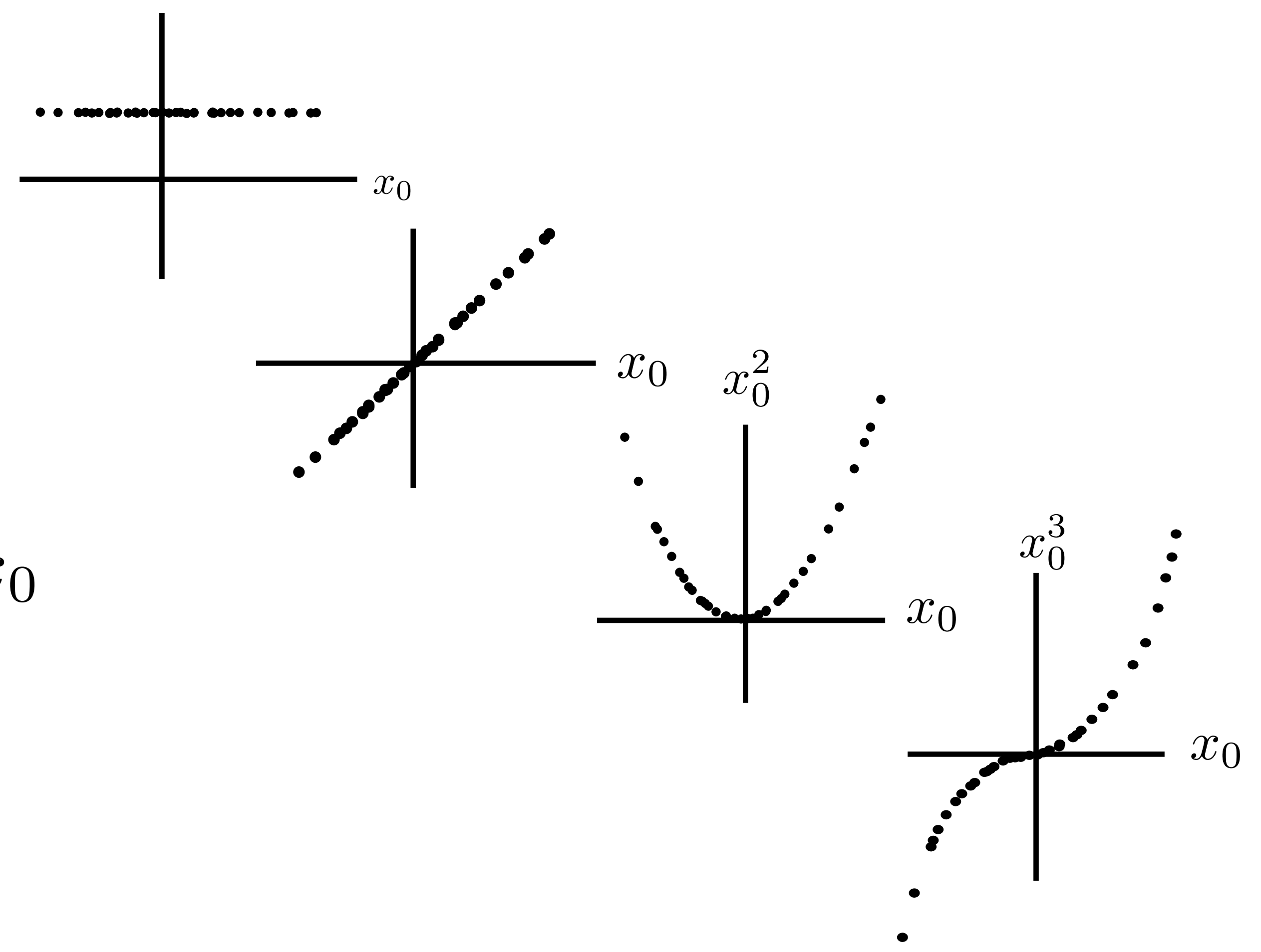
$$\sum_t \max\{0, y_t - f(h_t(x_t, \xi_t))\}$$

$$h_t(x_t) = [1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3]$$

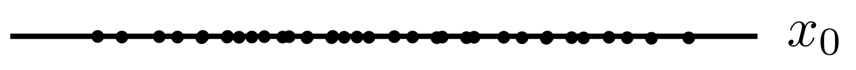
$$\max\{0, (\cdot)\}$$







$\Leftarrow$



Loss Functions

OUTPUTS
(Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} h_{00}(x_0,\xi_0) & \cdots & h_{0n}(x_0,\xi_0) \\ h_{10}(x_1,\xi_1) & \cdots & h_{1n}(x_1,\xi_1) \\ h_{20}(x_2,\xi_2) & \cdots & h_{2n}(x_2,\xi_2) \\ h_{30}(x_3,\xi_3) & \cdots & h_{3n}(x_3,\xi_3) \\ h_{40}(x_4,\xi_4) & \cdots & h_{4n}(x_4,\xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T,\xi_T) & \cdots & h_{Tn}(x_T,\xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

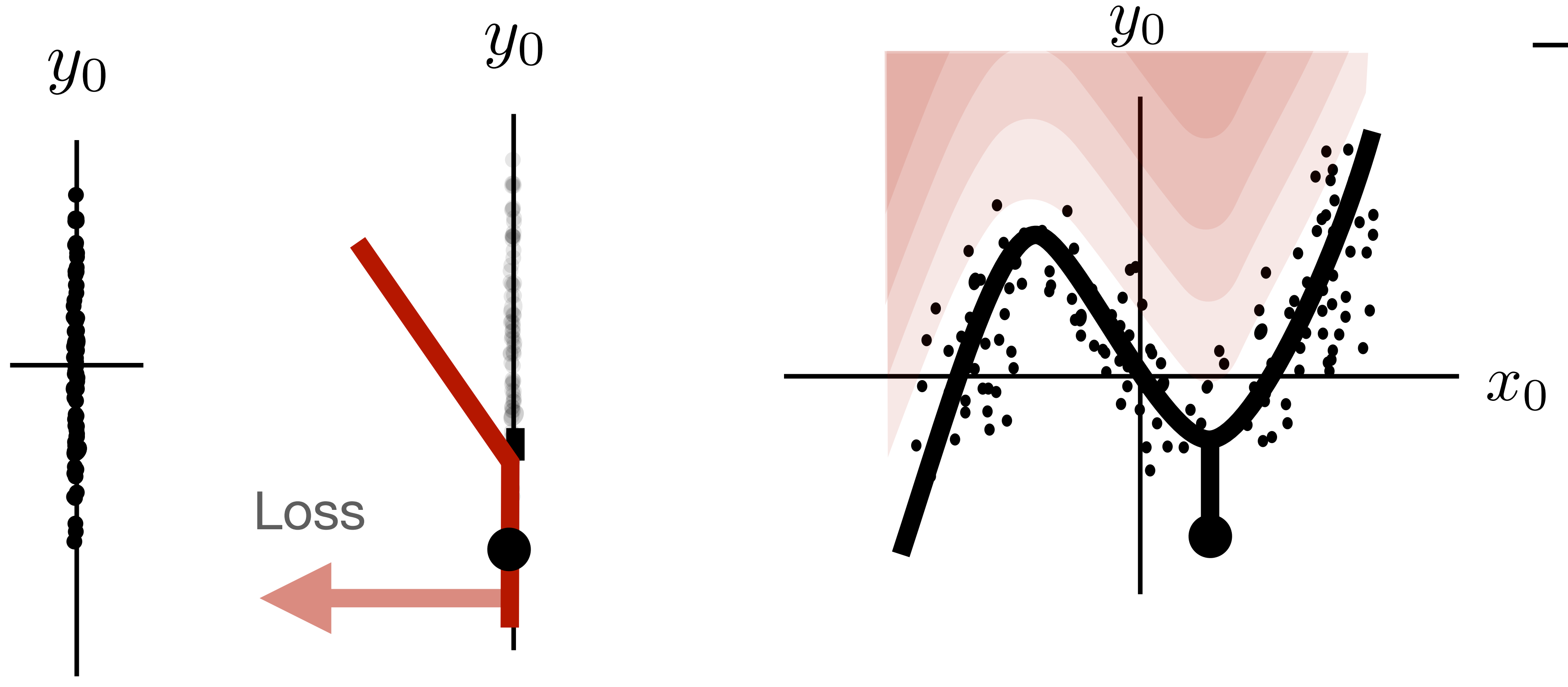
BASIS FUNCTIONS

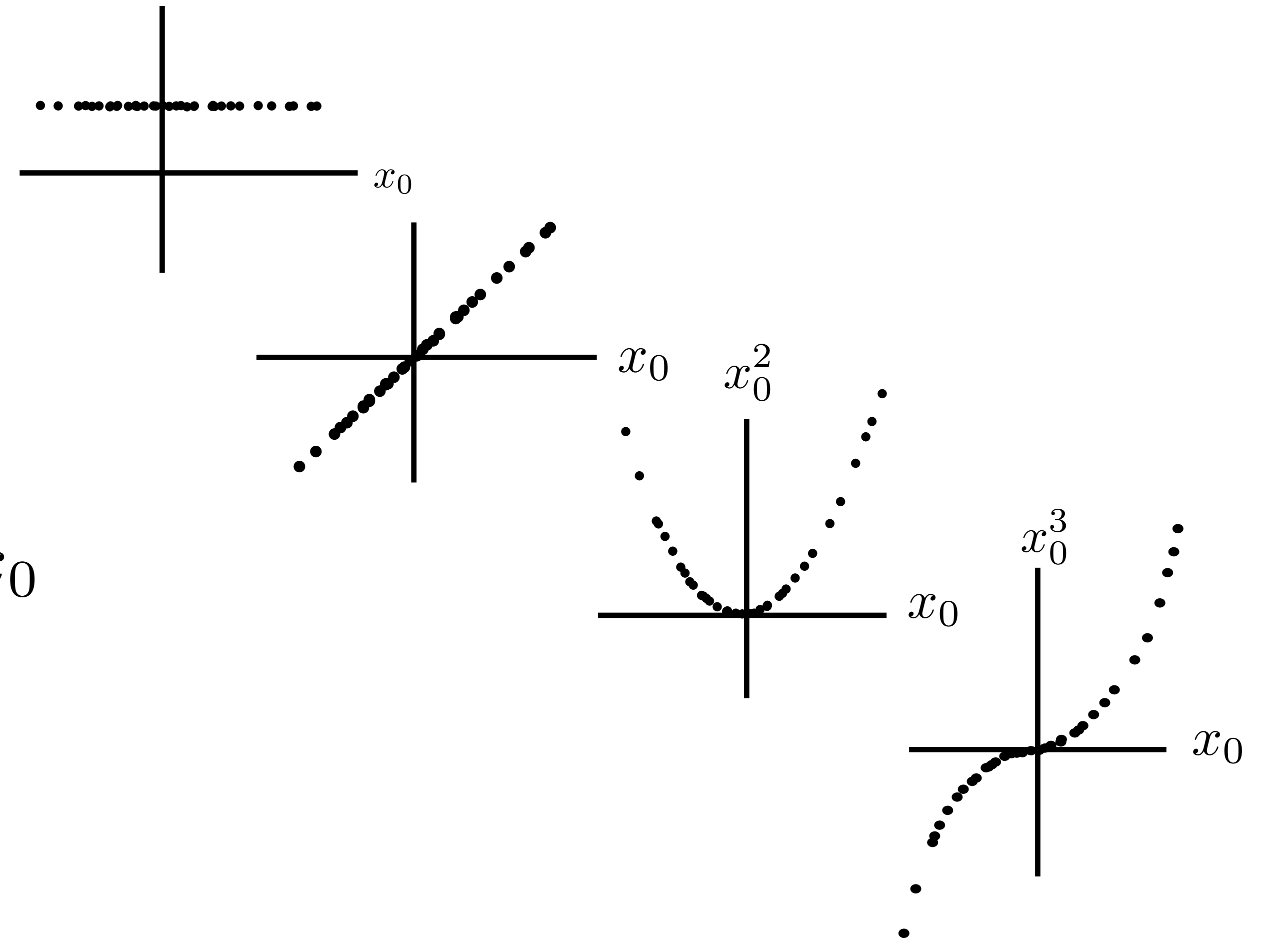
Hinge Loss

$$\sum_t \max\{0, y_t - f(h_t(x_t,\xi_t))\}$$

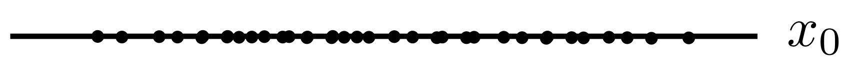
$$h_t(x_t) = [1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3]$$

$$\max\{0, (\cdot)\}$$





$\Leftarrow$



Loss Functions

OUTPUTS (Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS (Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0,\xi_0) & \cdots & h_{0n}(x_0,\xi_0) \\ h_{10}(x_1,\xi_1) & \cdots & h_{1n}(x_1,\xi_1) \\ h_{20}(x_2,\xi_2) & \cdots & h_{2n}(x_2,\xi_2) \\ h_{30}(x_3,\xi_3) & \cdots & h_{3n}(x_3,\xi_3) \\ h_{40}(x_4,\xi_4) & \cdots & h_{4n}(x_4,\xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T,\xi_T) & \cdots & h_{Tn}(x_T,\xi_T) \end{bmatrix} \Longleftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

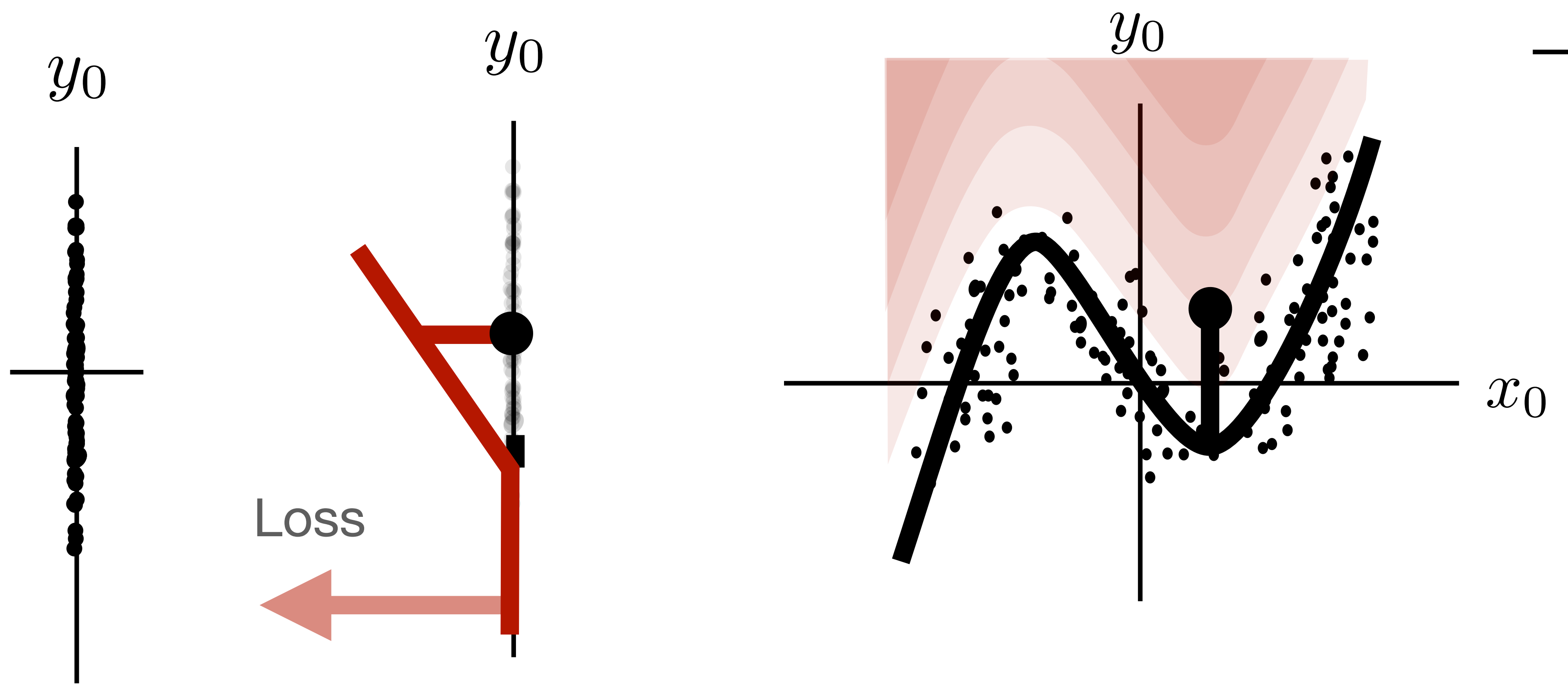
BASIS FUNCTIONS

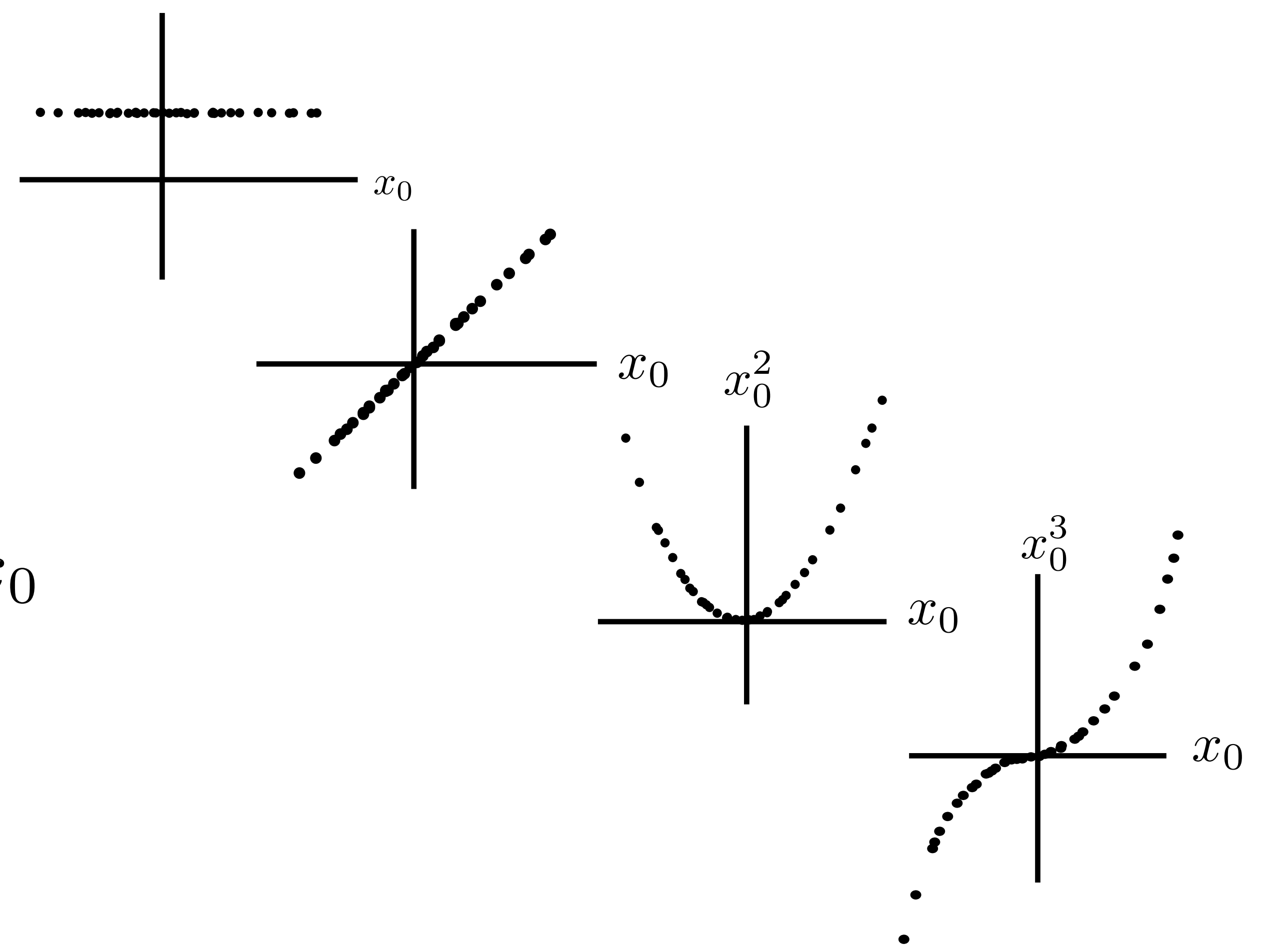
Hinge Loss

$$\sum_t \max\{0, y_t - f(h_t(x_t, \xi_t))\}$$

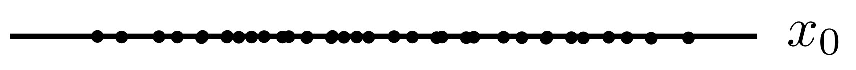
$$h_t(x_t) = [1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3]$$

$$\max\{0, (\cdot)\}$$





$\Longleftarrow$



Loss Functions

OUTPUTS
(Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

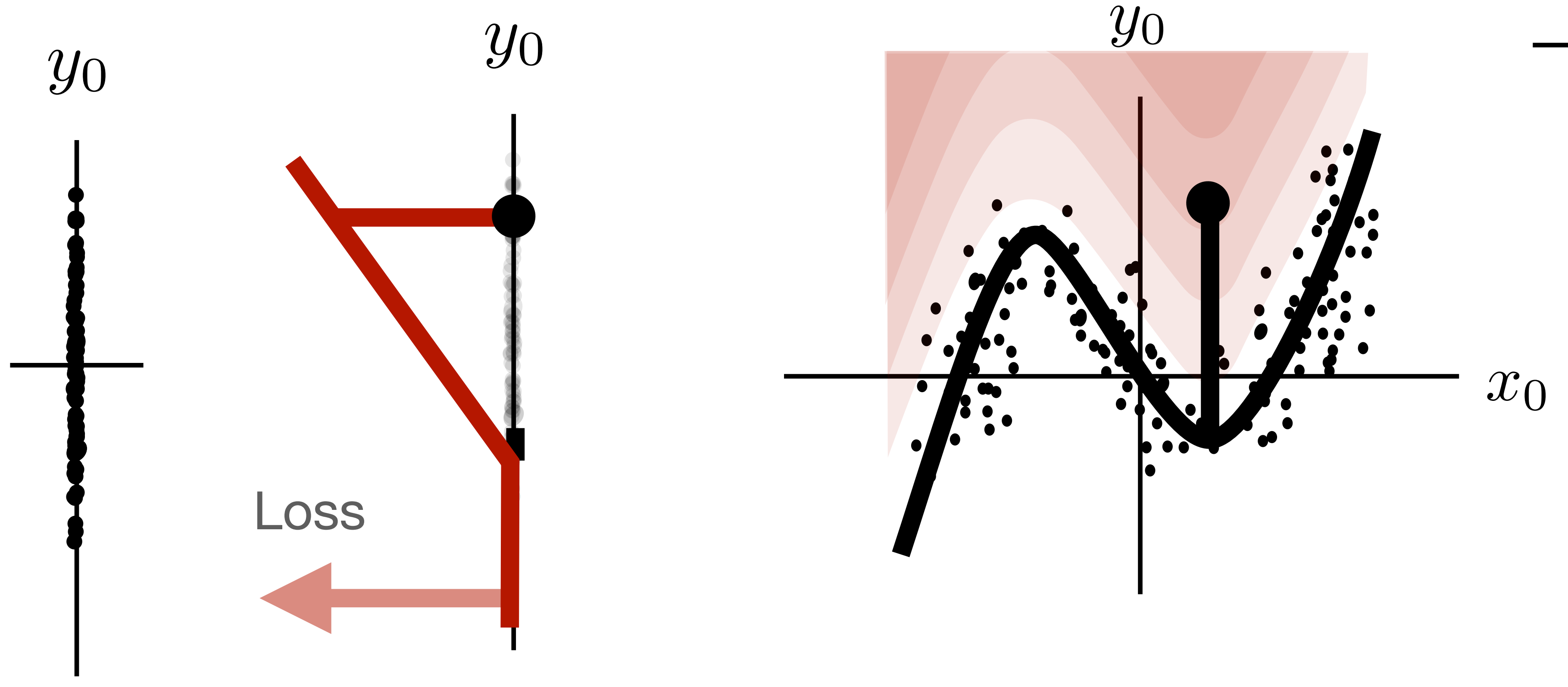
BASIS FUNCTIONS

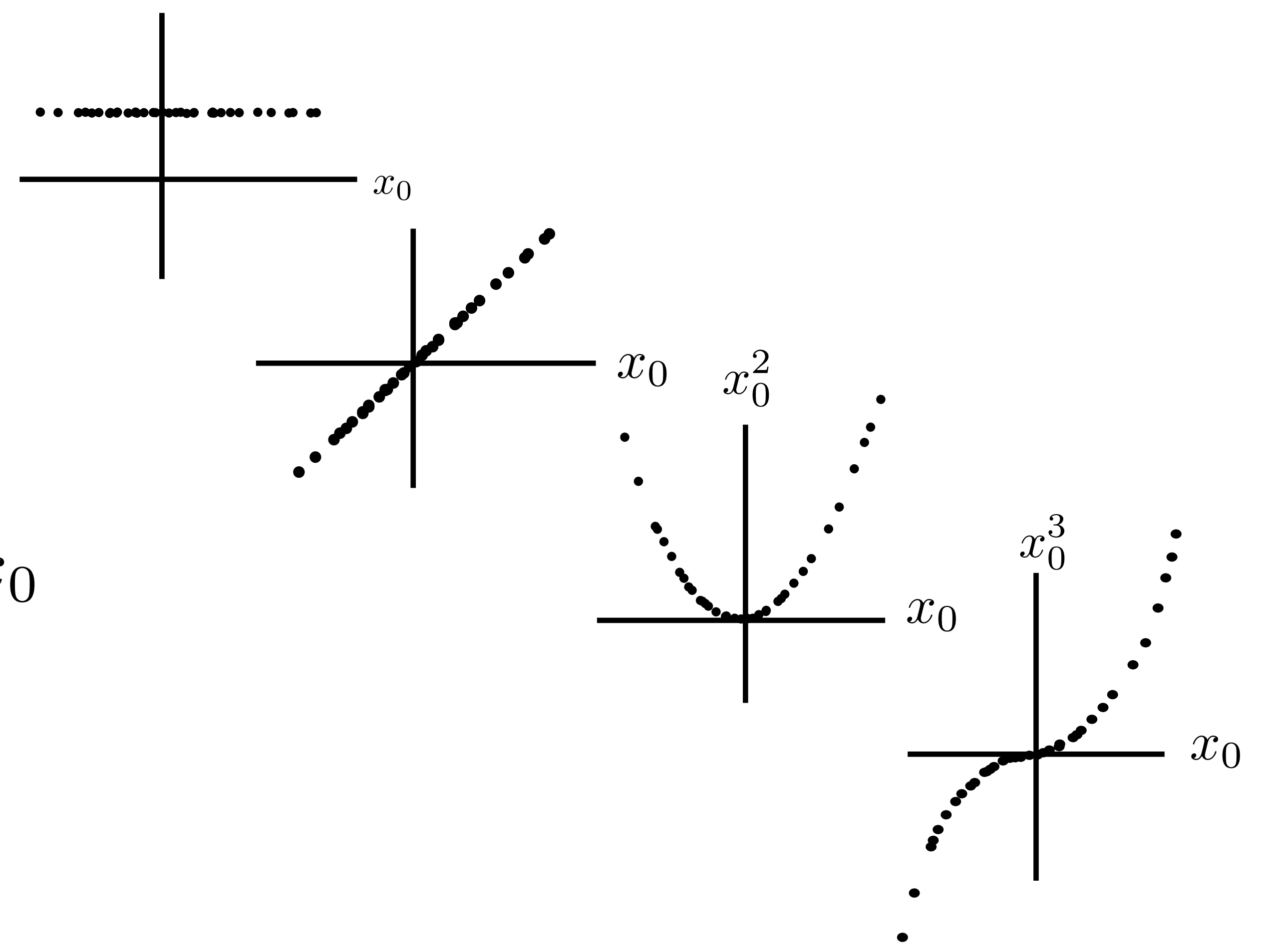
Hinge Loss

$$\sum_t \max\{0, y_t - f(h_t(x_t, \xi_t))\}$$

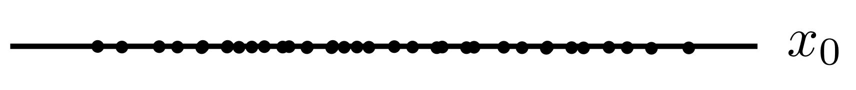
$$h_t(x_t) = [1 \ x_{t0} \ x_{t0}^2 \ x_{t0}^3]$$

$$\max\{0, (\cdot)\}$$





$\Leftarrow$



Loss Functions

OUTPUTS (Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS (Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

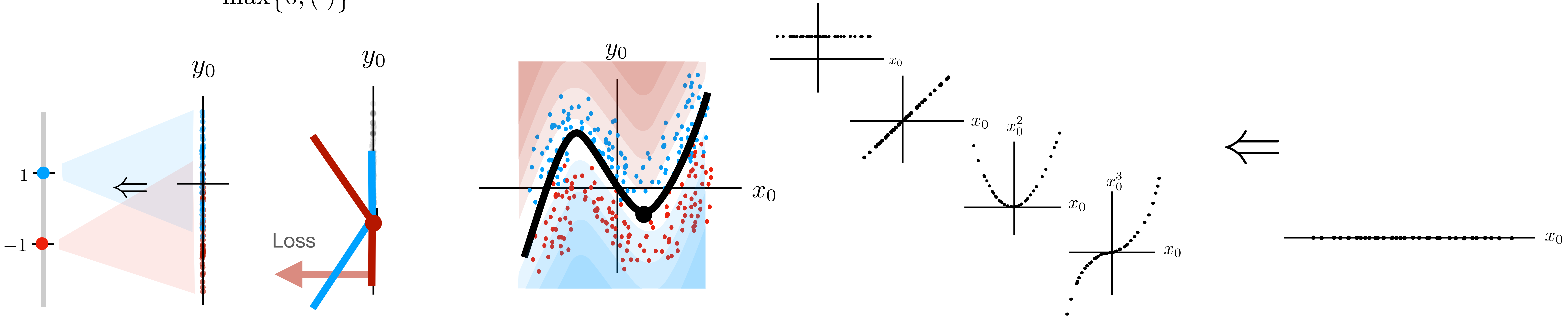
BASIS FUNCTIONS

Hinge Loss

$$\sum_t \max\left\{0, \gamma_t\left(y_t - f(h_t(x_t, \xi_t))\right)\right\}$$

$$h_t(x_t) = [1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3]$$

$$\max\{0, (\cdot)\}$$



Loss Functions

OUTPUTS (Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS (Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

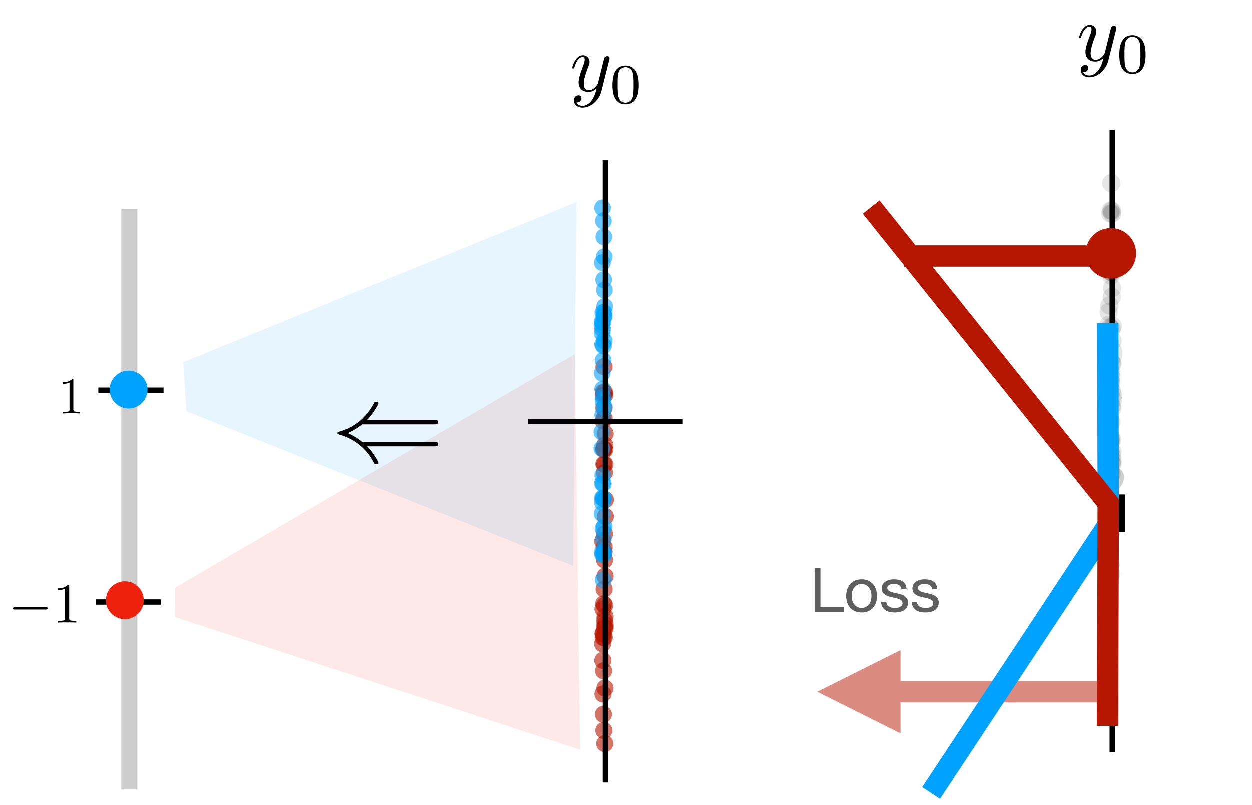
BASIS FUNCTIONS

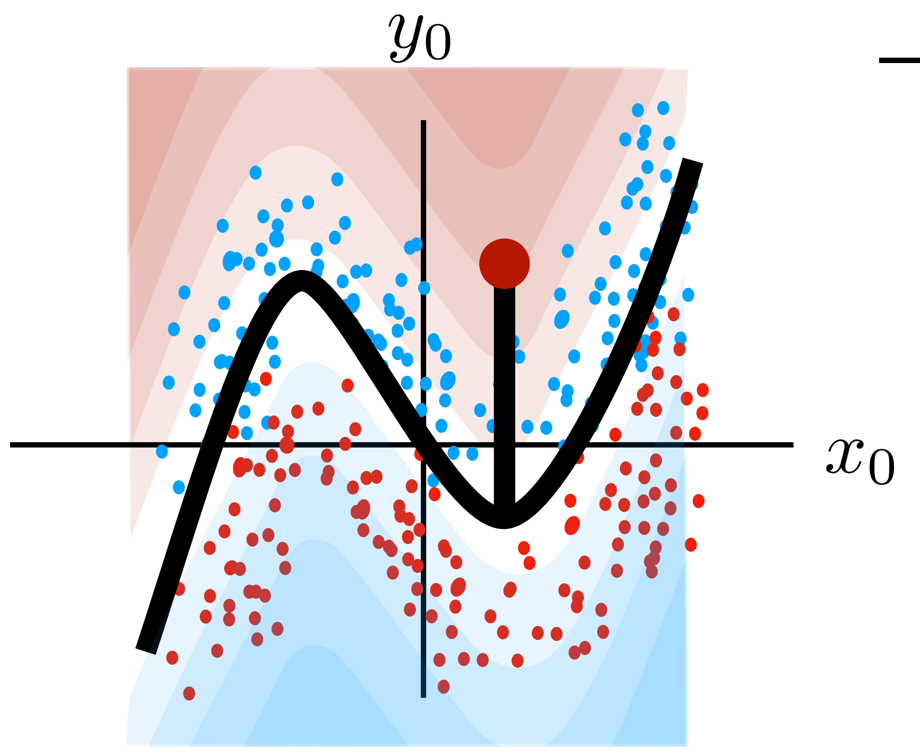
Hinge Loss

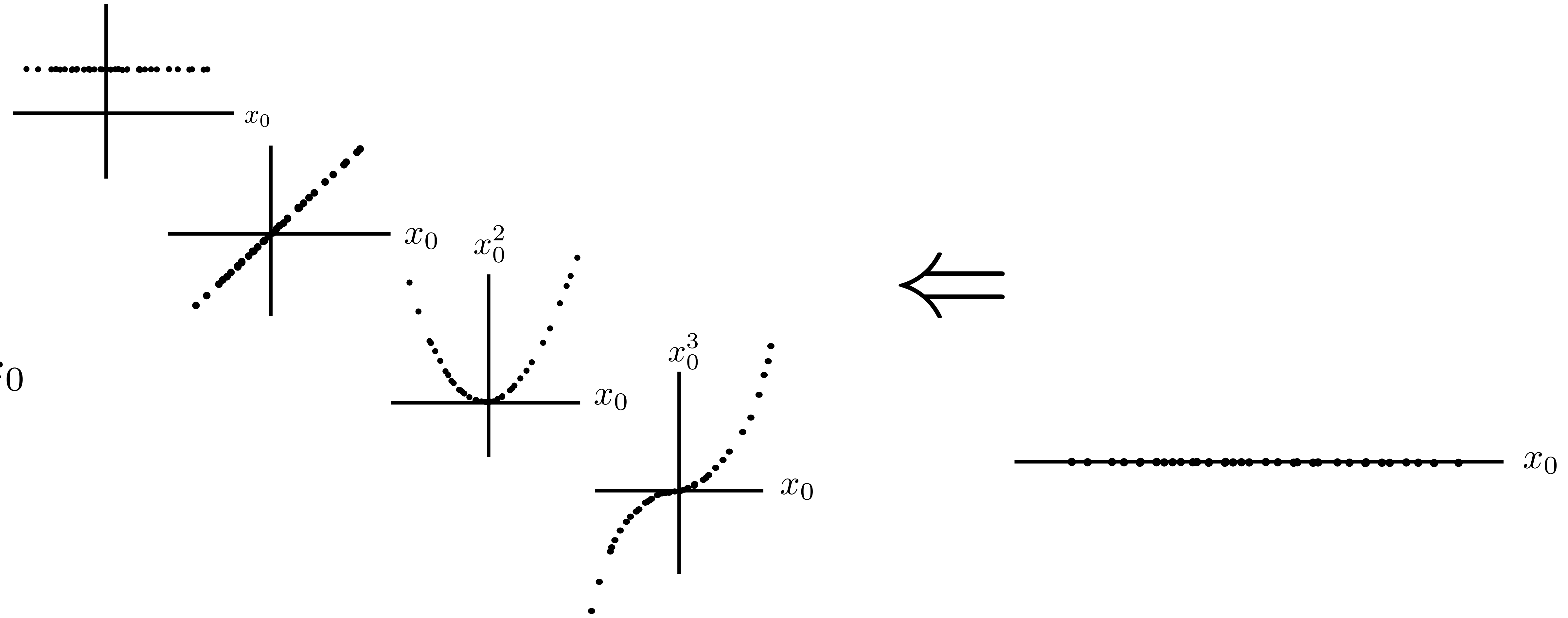
$$\sum_t \max\left\{0, \gamma_t(y_t - f(h_t(x_t, \xi_t)))\right\}$$

$$h_t(x_t) = [1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3]$$

$$\max\{0, (\cdot)\}$$







Loss Functions

OUTPUTS (Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS (Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Longleftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

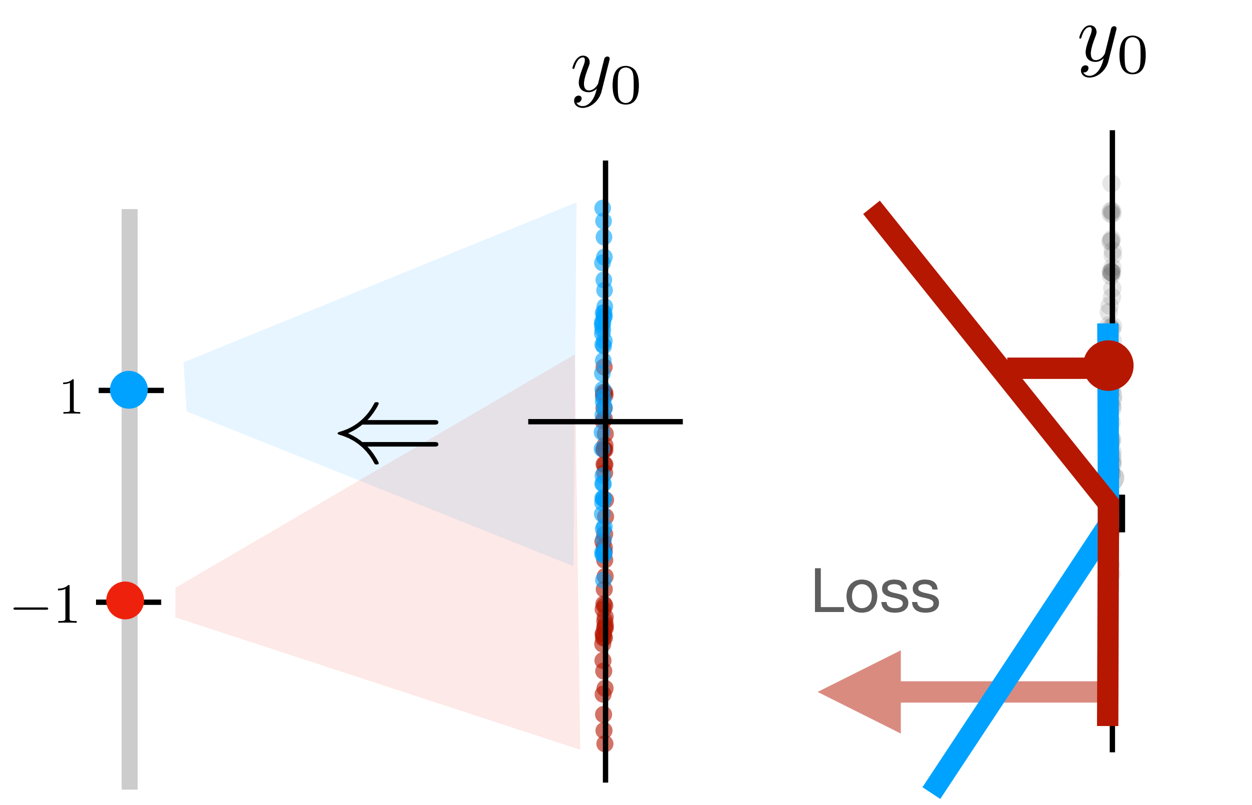
BASIS FUNCTIONS

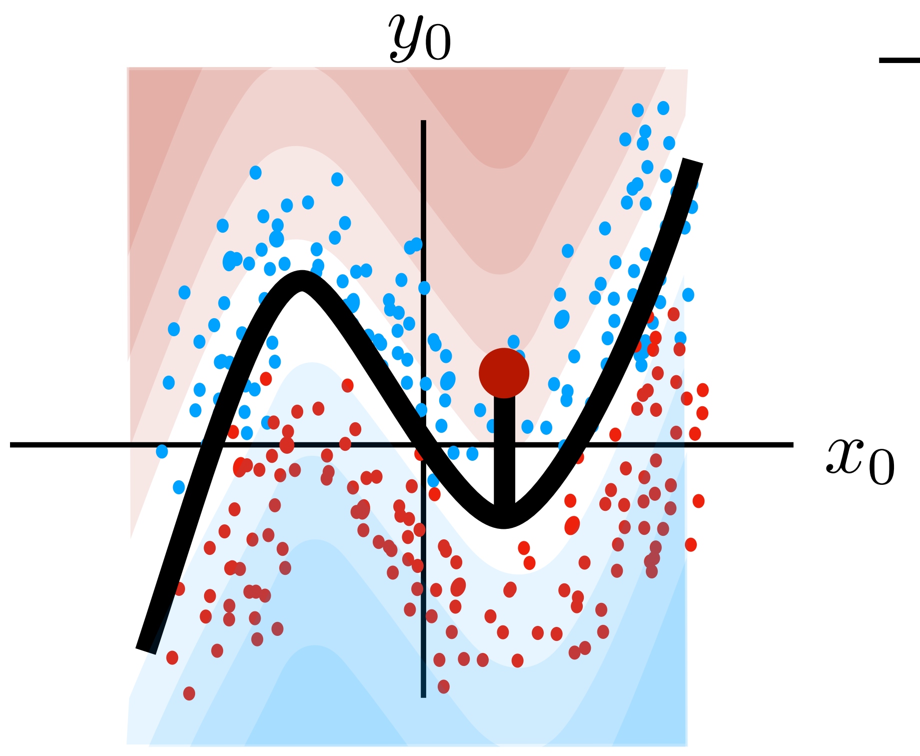
Hinge Loss

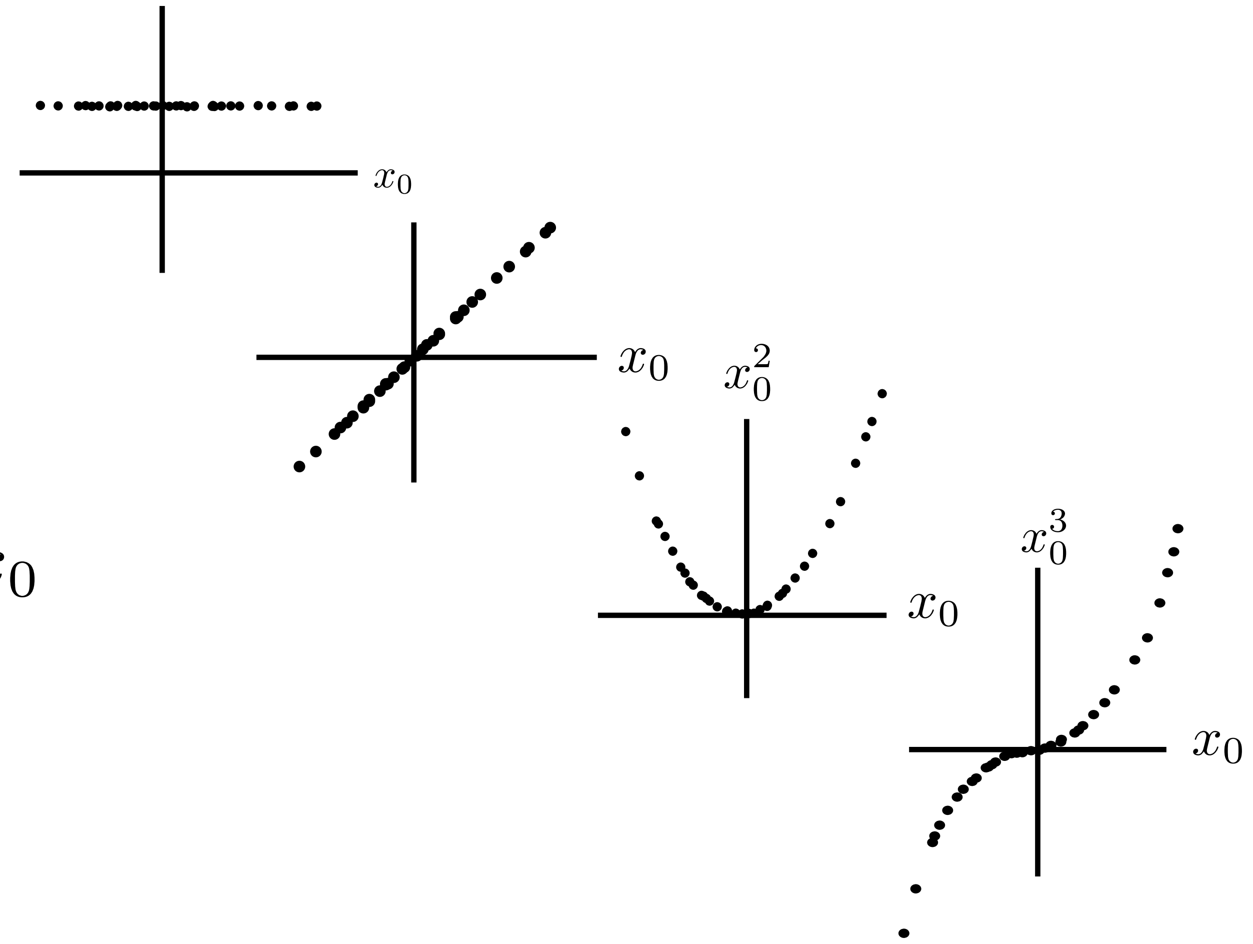
$$\sum_t \max\{0, \gamma_t(y_t - f(h_t(x_t, \xi_t)))\}$$

$$h_t(x_t) = [1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3]$$

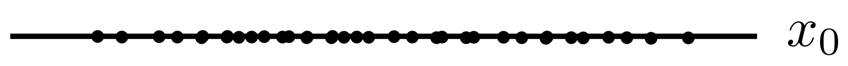
$$\max\{0, (\cdot)\}$$







$\Longleftarrow$



Loss Functions

OUTPUTS (Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS (Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} h_{00}(x_0,\xi_0) & \cdots & h_{0n}(x_0,\xi_0) \\ h_{10}(x_1,\xi_1) & \cdots & h_{1n}(x_1,\xi_1) \\ h_{20}(x_2,\xi_2) & \cdots & h_{2n}(x_2,\xi_2) \\ h_{30}(x_3,\xi_3) & \cdots & h_{3n}(x_3,\xi_3) \\ h_{40}(x_4,\xi_4) & \cdots & h_{4n}(x_4,\xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T,\xi_T) & \cdots & h_{Tn}(x_T,\xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

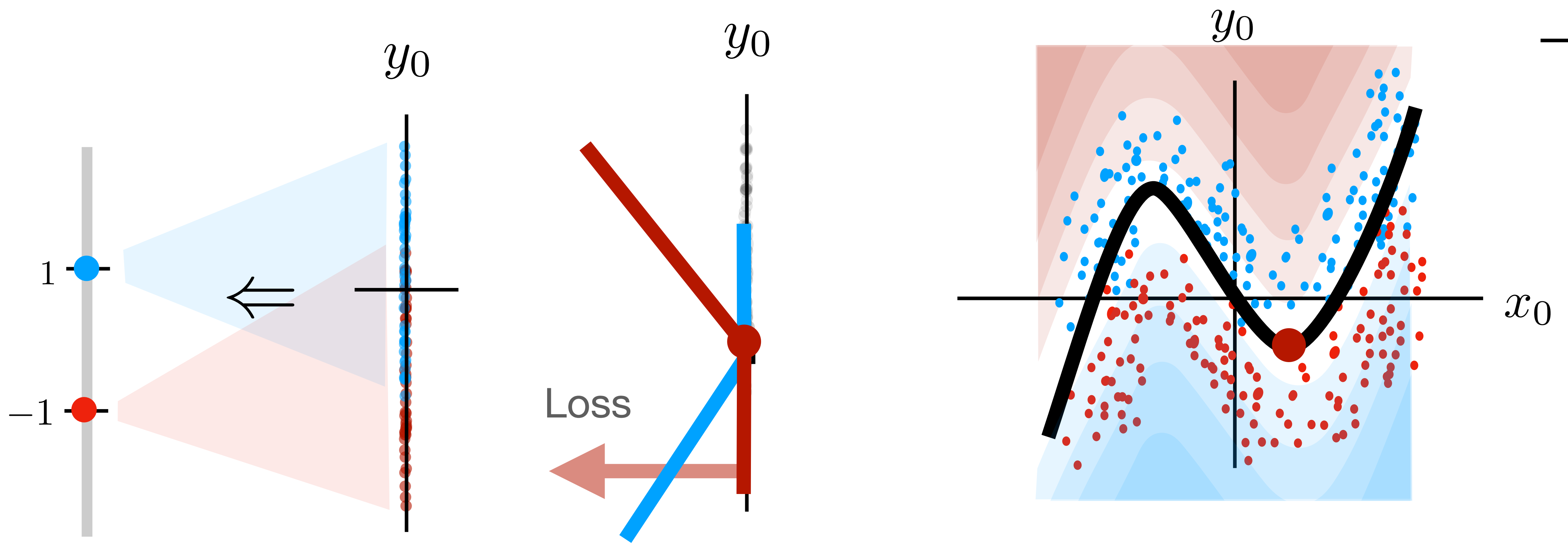
BASIS FUNCTIONS

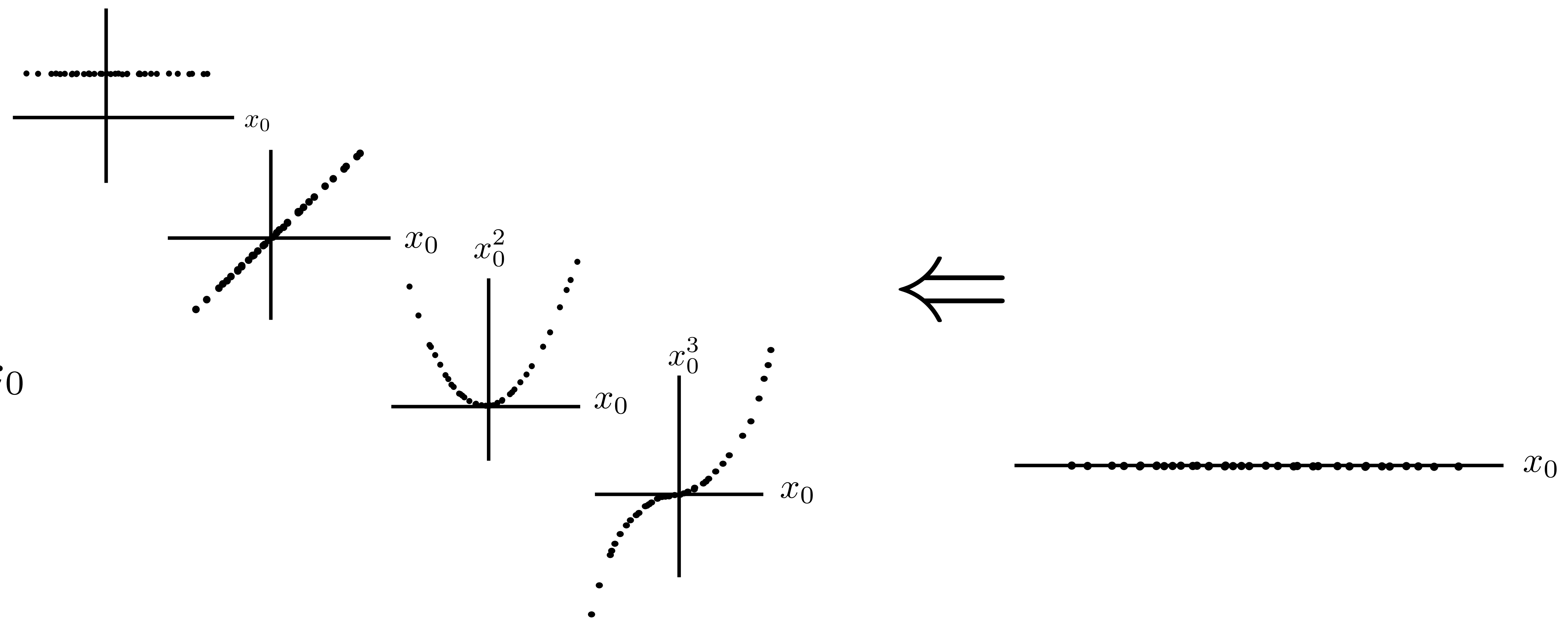
Hinge Loss

$$\sum_t \max\{0, \gamma_t(y_t - f(h_t(x_t, \xi_t)))\}$$

$$h_t(x_t) = [1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3]$$

$$\max\{0, (\cdot)\}$$





Loss Functions

OUTPUTS (Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS (Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0,\xi_0) & \cdots & h_{0n}(x_0,\xi_0) \\ h_{10}(x_1,\xi_1) & \cdots & h_{1n}(x_1,\xi_1) \\ h_{20}(x_2,\xi_2) & \cdots & h_{2n}(x_2,\xi_2) \\ h_{30}(x_3,\xi_3) & \cdots & h_{3n}(x_3,\xi_3) \\ h_{40}(x_4,\xi_4) & \cdots & h_{4n}(x_4,\xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T,\xi_T) & \cdots & h_{Tn}(x_T,\xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

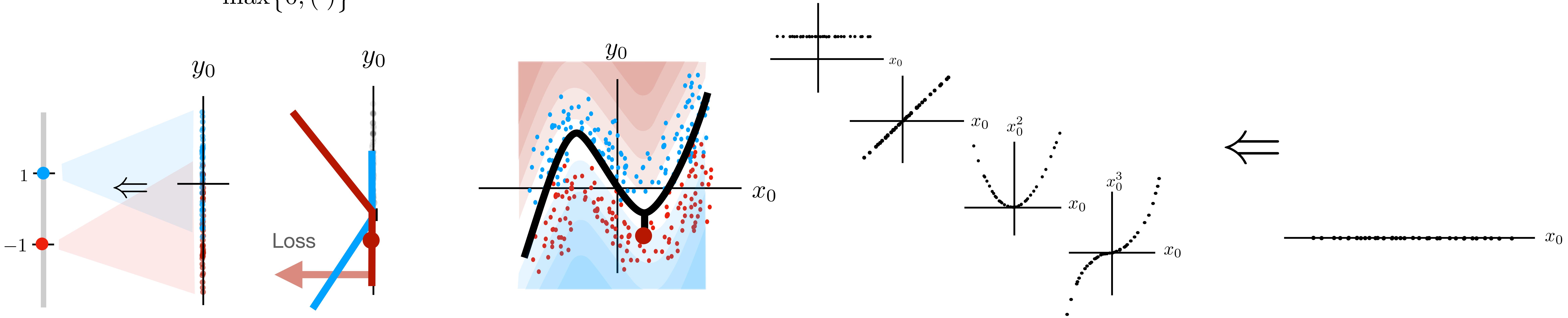
BASIS FUNCTIONS

Hinge Loss

$$\sum_t \max\left\{0, \gamma_t\left(y_t - f(h_t(x_t, \xi_t))\right)\right\}$$

$$h_t(x_t) = [1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3]$$

$$\max\{0, (\cdot)\}$$



Loss Functions

OUTPUTS
(Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

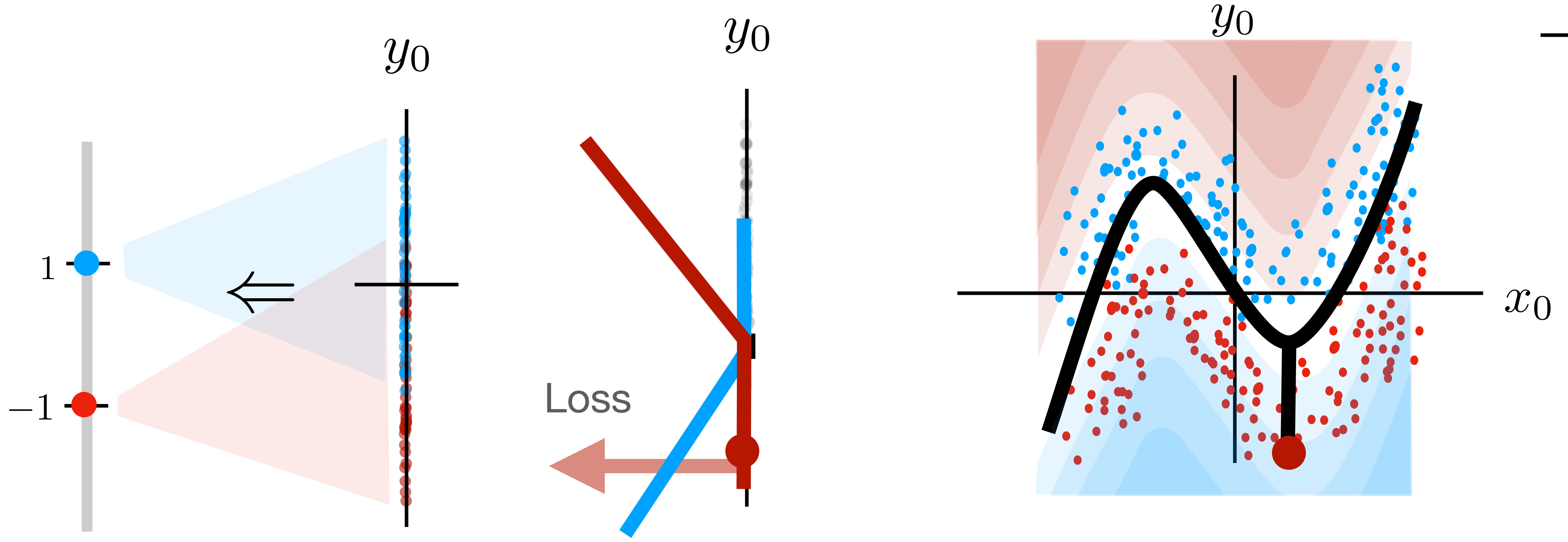
BASIS FUNCTIONS

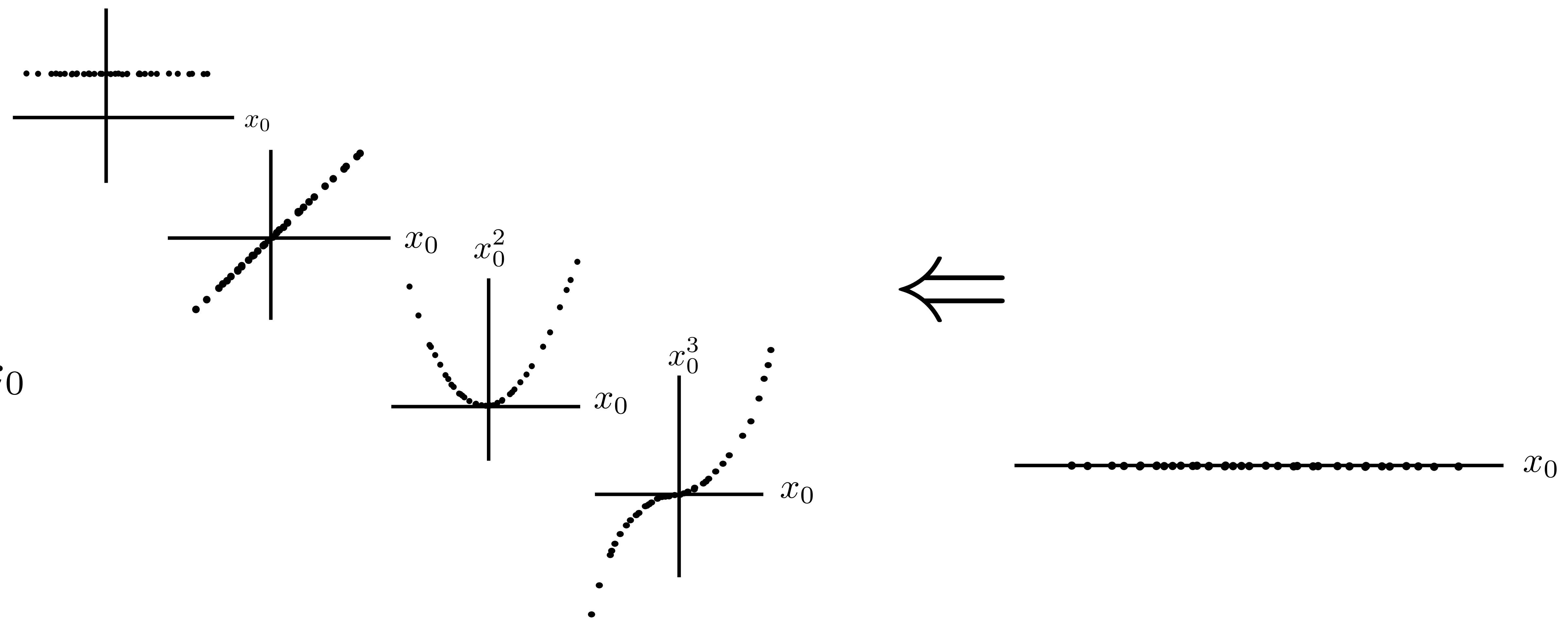
Hinge Loss

$$\sum_t \max\left\{0, \gamma_t\left(y_t - f(h_t(x_t, \xi_t))\right)\right\}$$

$$h_t(x_t) = \left[1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3\right]$$

$$\max\{0, (\cdot)\}$$





Loss Functions

OUTPUTS (Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS (Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

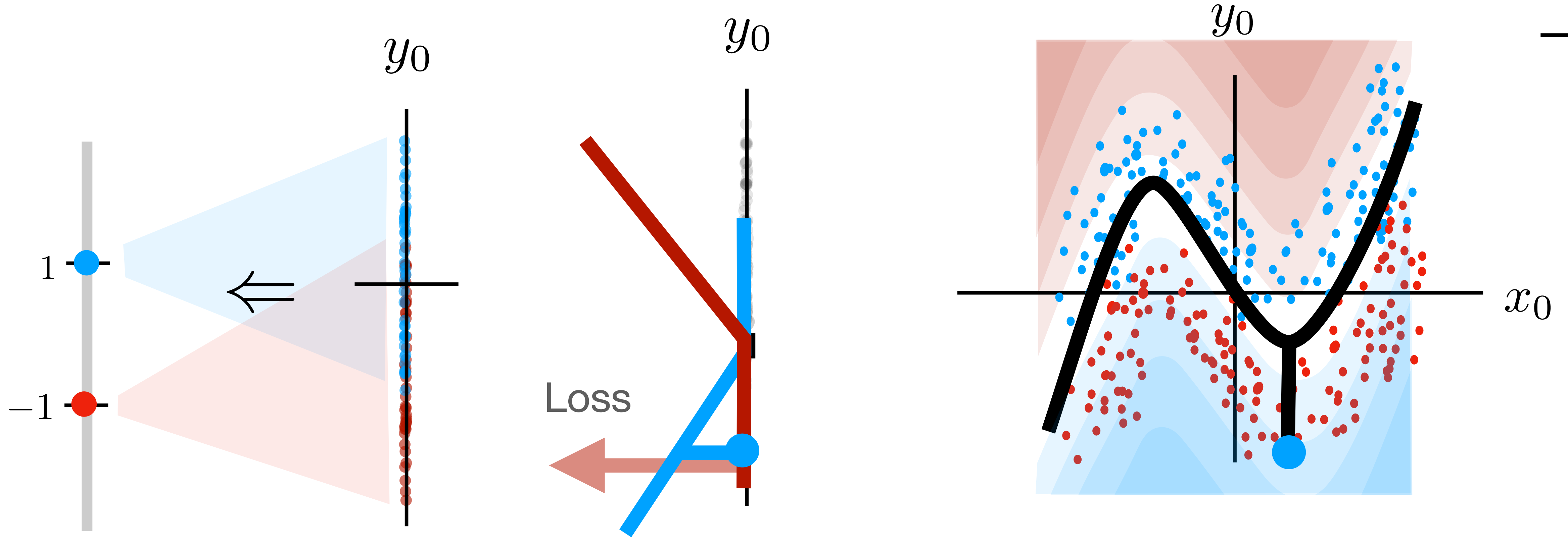
BASIS FUNCTIONS

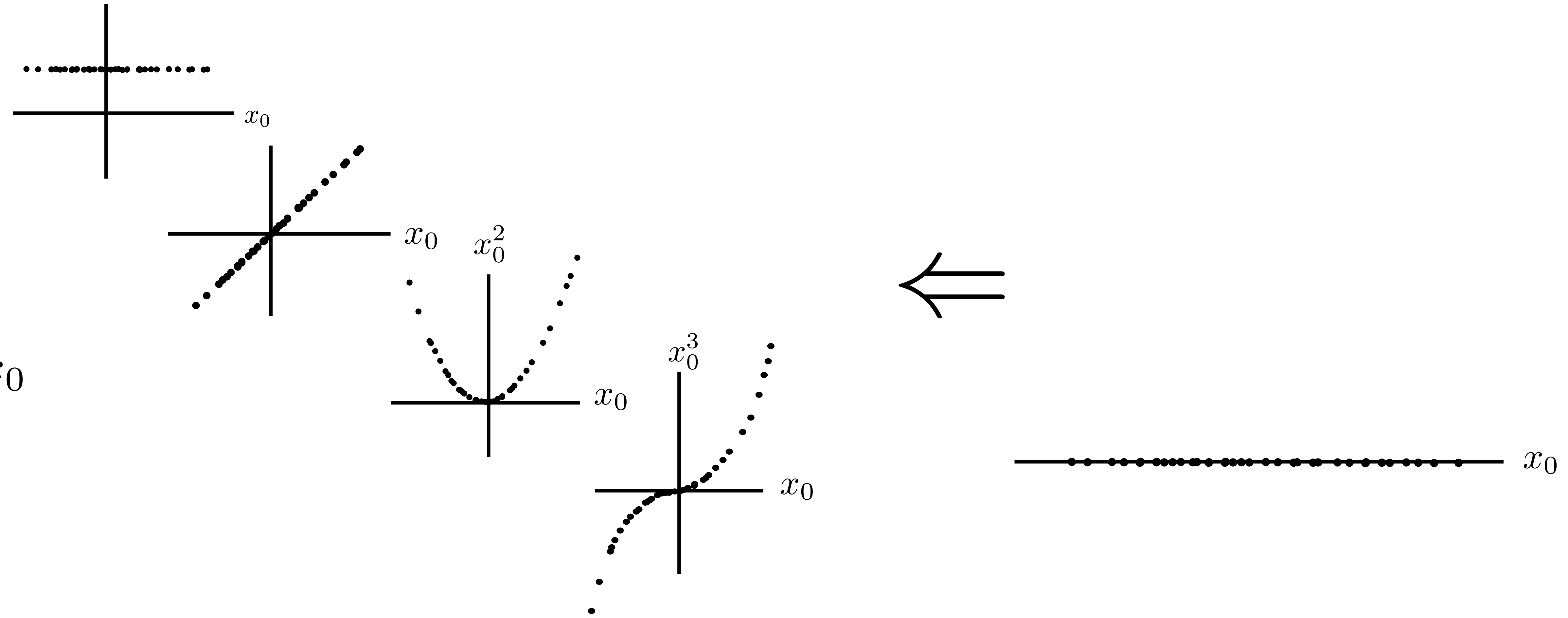
Hinge Loss

$$\sum_t \max\left\{0, \gamma_t\left(y_t - f(h_t(x_t, \xi_t))\right)\right\}$$

$$h_t(x_t) = [1 \ x_{t0} \ x_{t0}^2 \ x_{t0}^3]$$

$$\max\{0, (\cdot)\}$$





Loss Functions

OUTPUTS (Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS (Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Longleftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

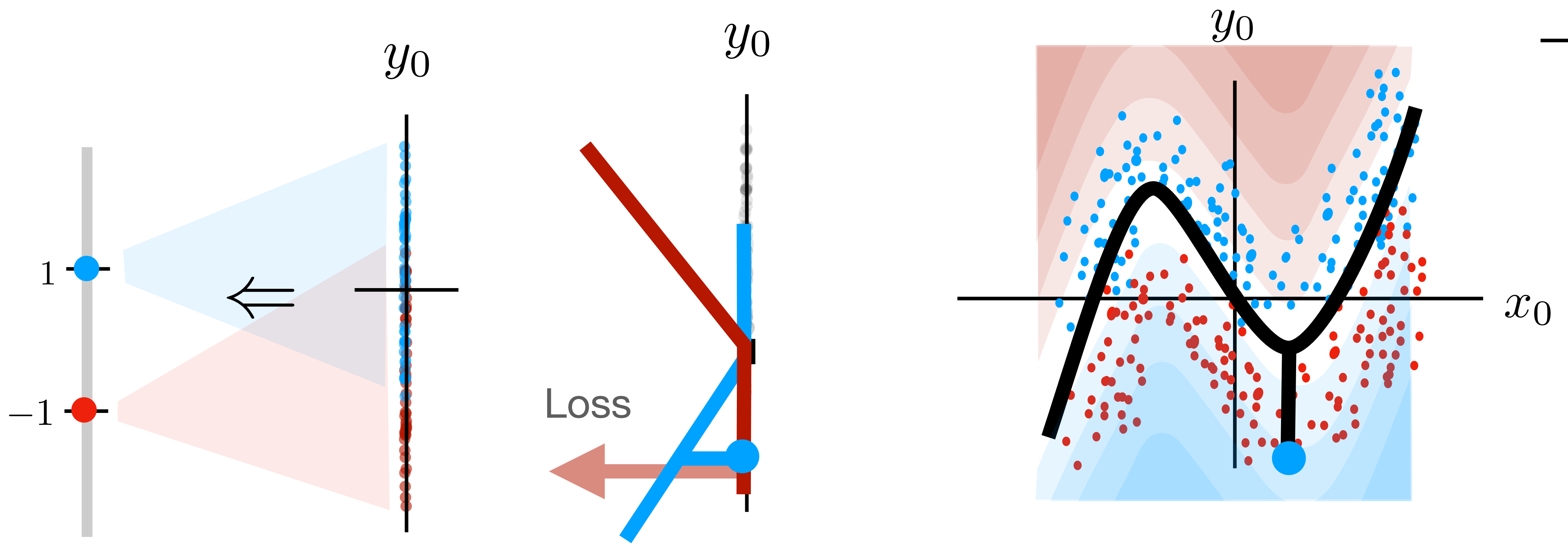
BASIS FUNCTIONS

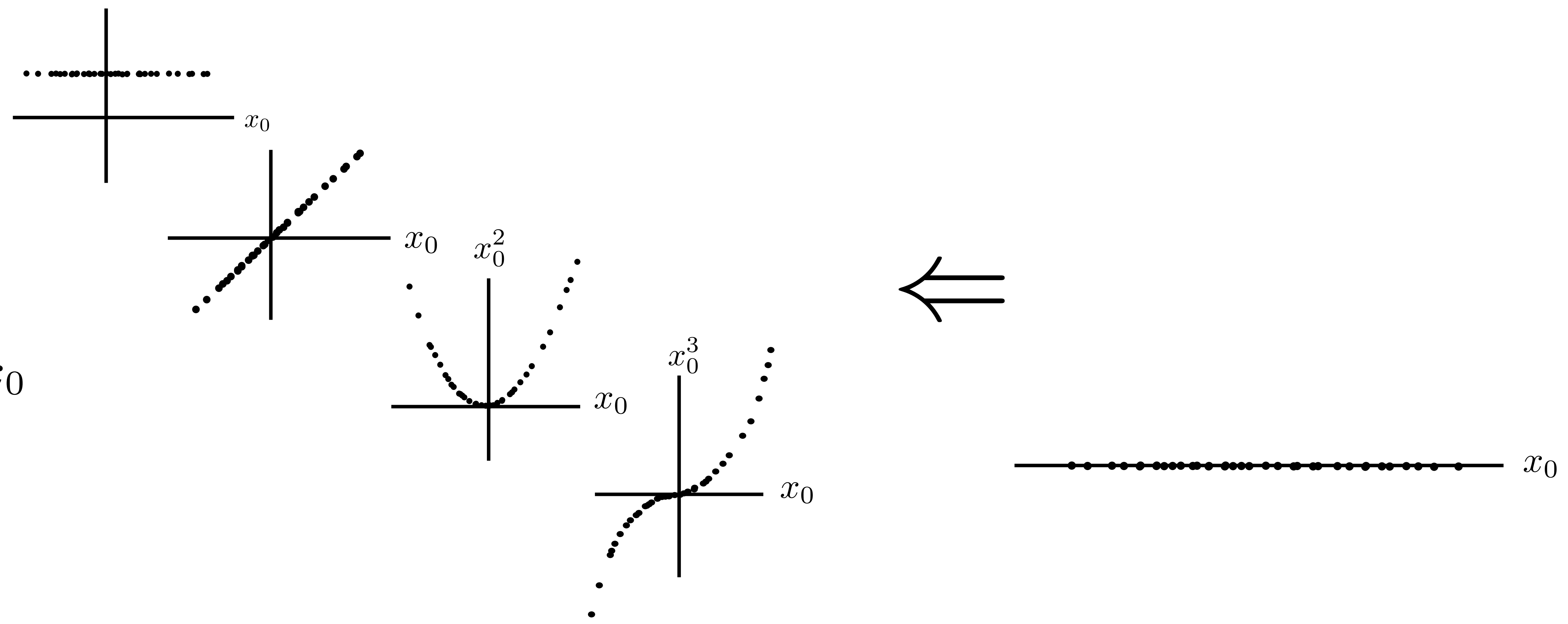
Hinge Loss

$$\sum_t \max\left\{0, \gamma_t\left(y_t - f(h_t(x_t, \xi_t))\right)\right\}$$

$$h_t(x_t) = [1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3]$$

$\max\{0, (\cdot)\}$





Loss Functions

OUTPUTS (Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS (Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

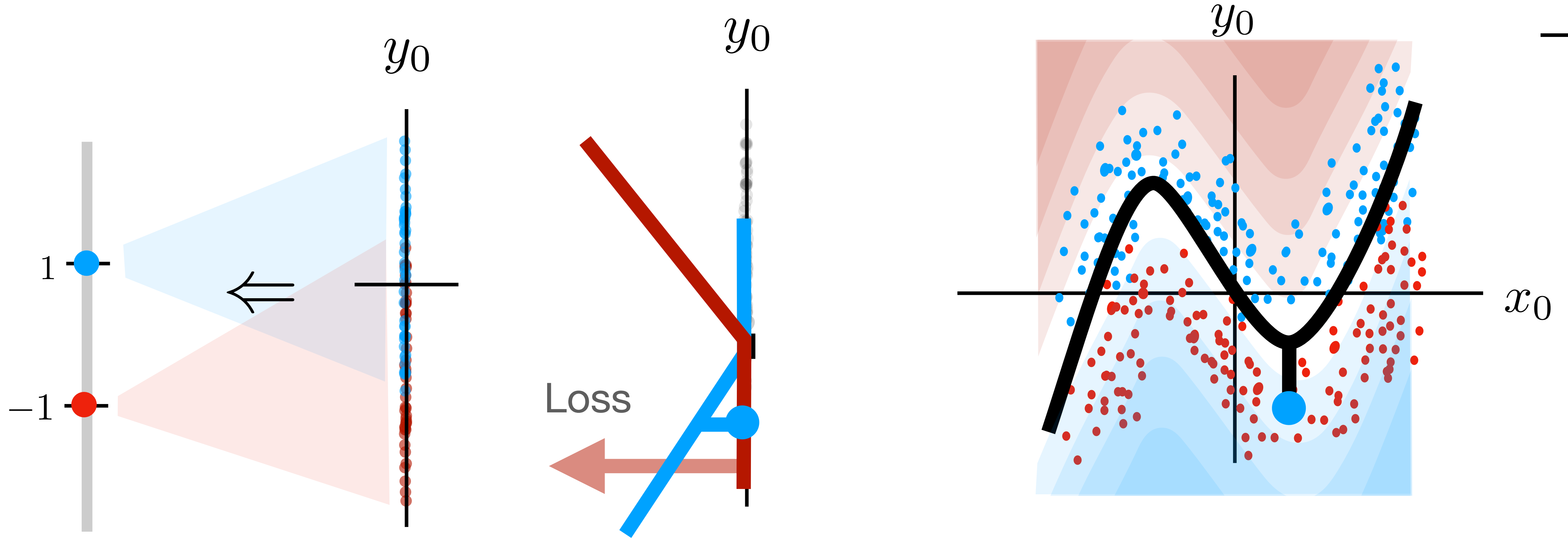
BASIS FUNCTIONS

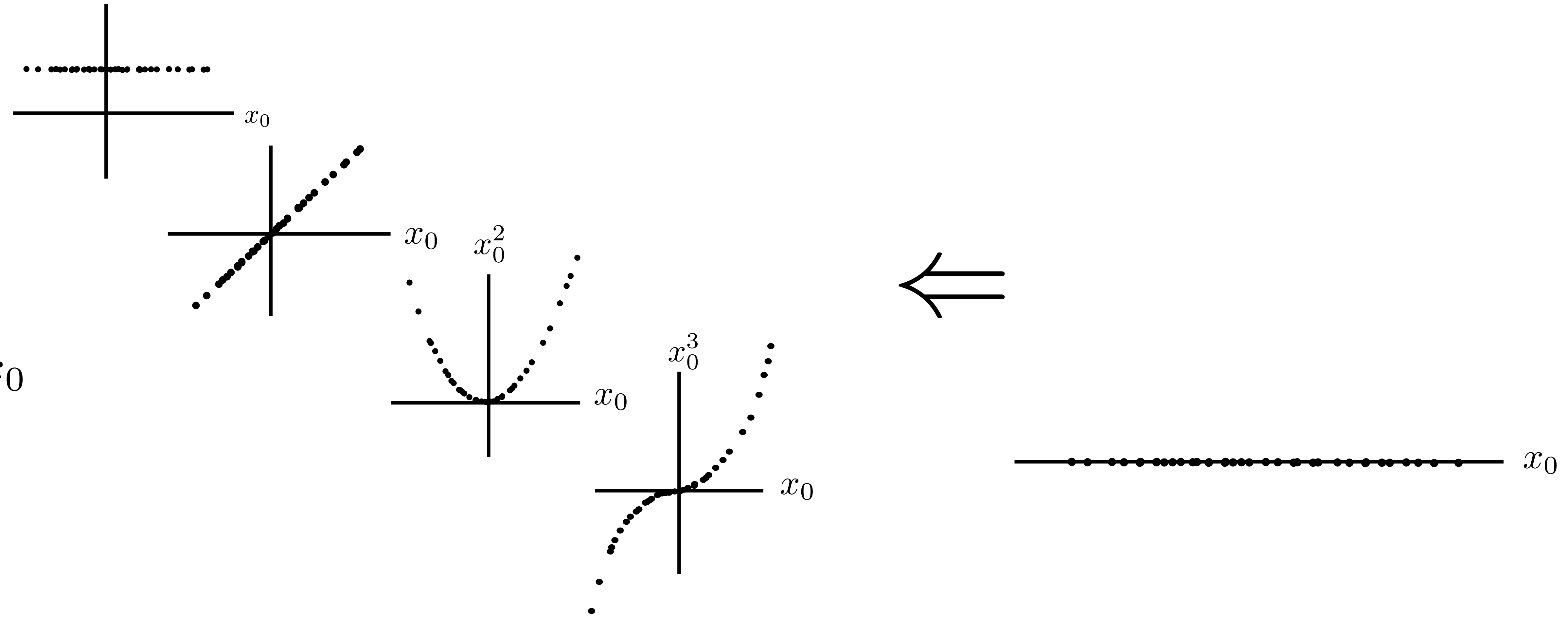
Hinge Loss

$$\sum_t \max\{0, \gamma_t(y_t - f(h_t(x_t, \xi_t)))\}$$

$$h_t(x_t) = [1 \ x_{t0} \ x_{t0}^2 \ x_{t0}^3]$$

$$\max\{0, (\cdot)\}$$





Loss Functions

OUTPUTS (Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS (Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

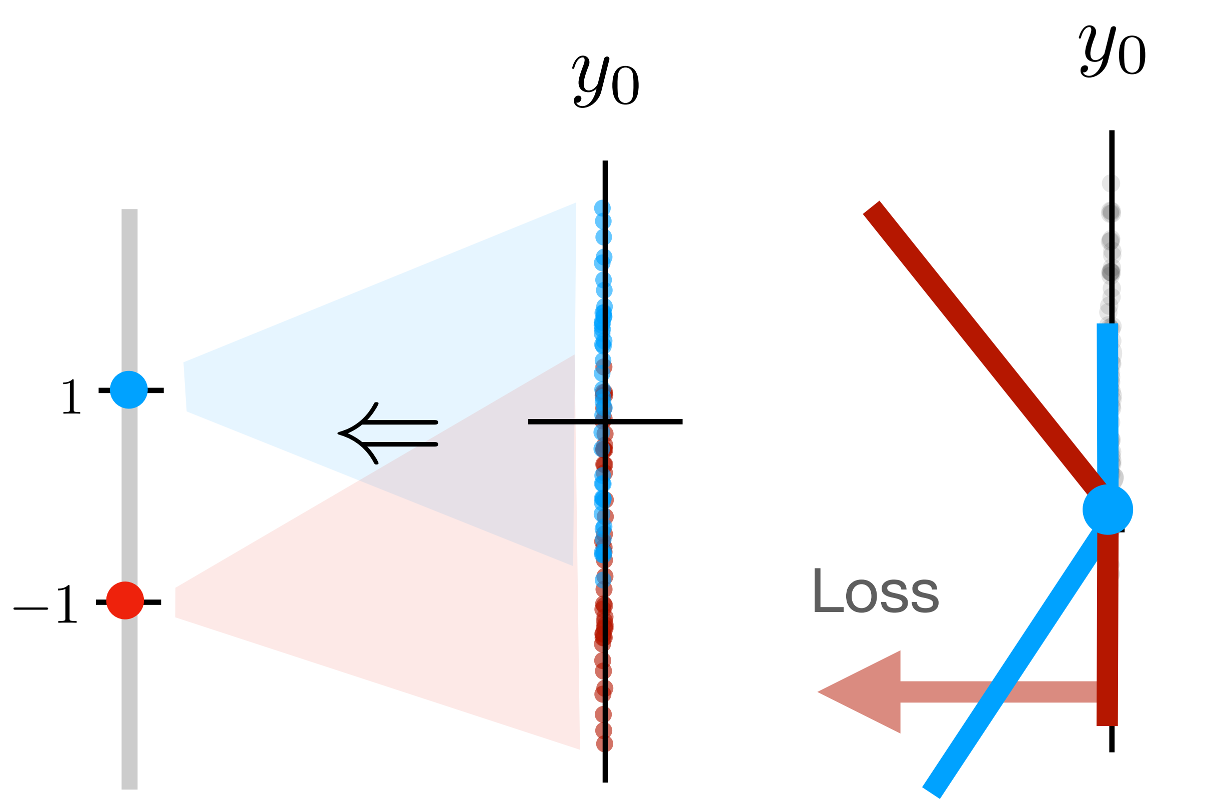
BASIS FUNCTIONS

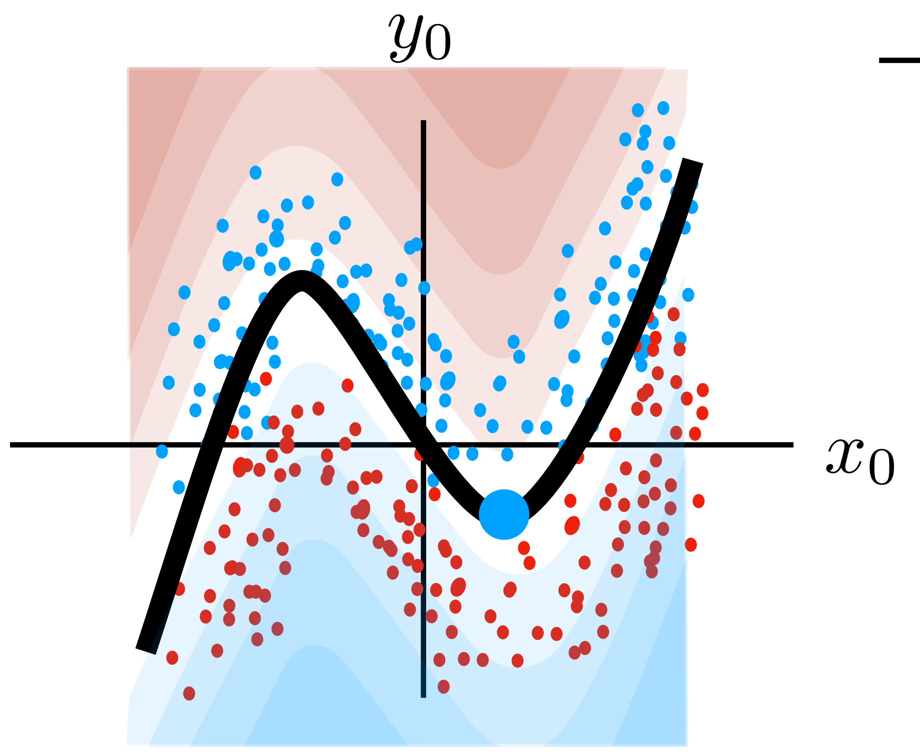
Hinge Loss

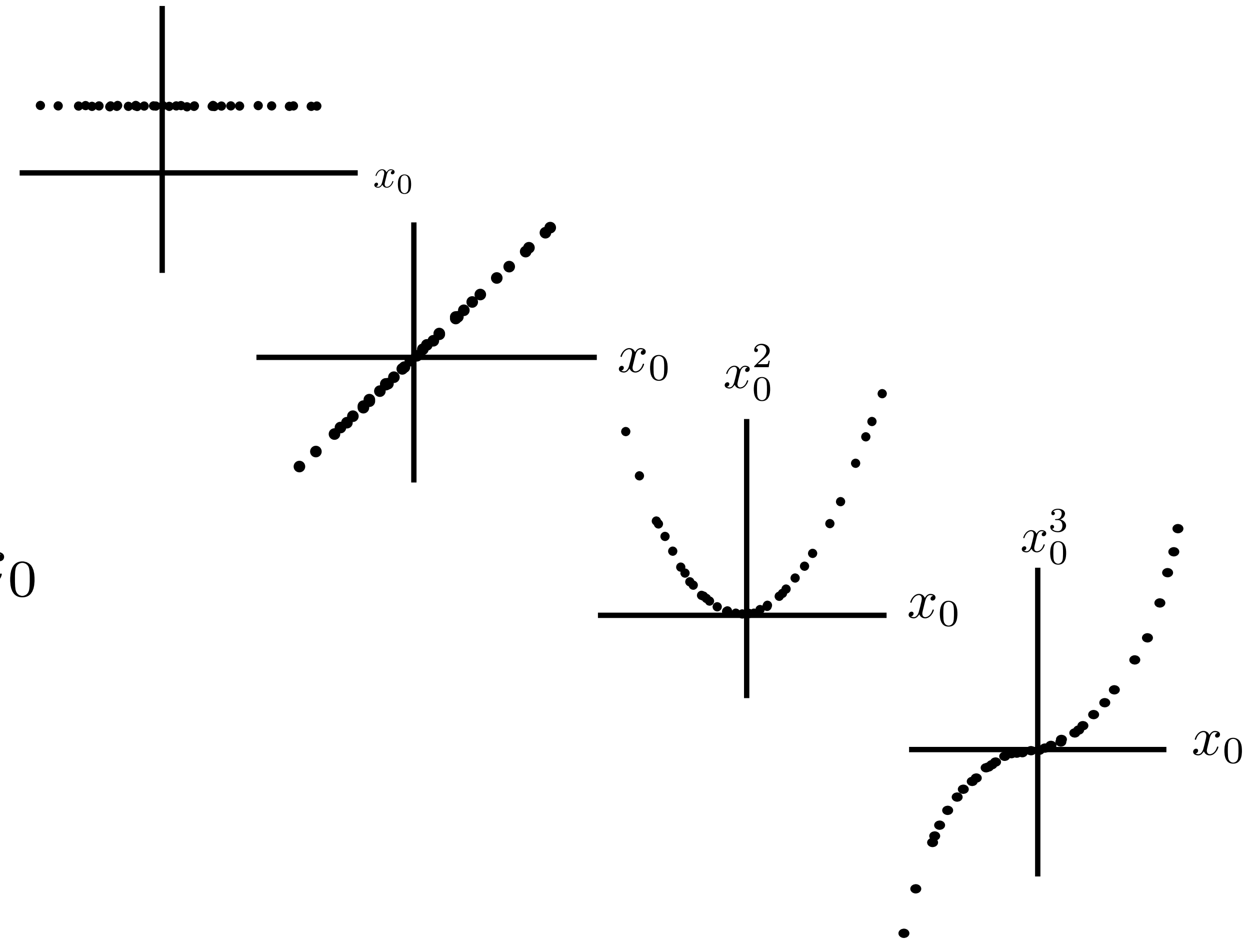
$$\sum_t \max\{0, \gamma_t(y_t - f(h_t(x_t, \xi_t)))\}$$

$$h_t(x_t) = [1 \ x_{t0} \ x_{t0}^2 \ x_{t0}^3]$$

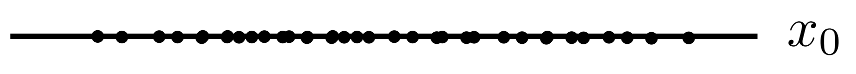
$$\max\{0, (\cdot)\}$$







$\Leftarrow$



Loss Functions

OUTPUTS
(Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

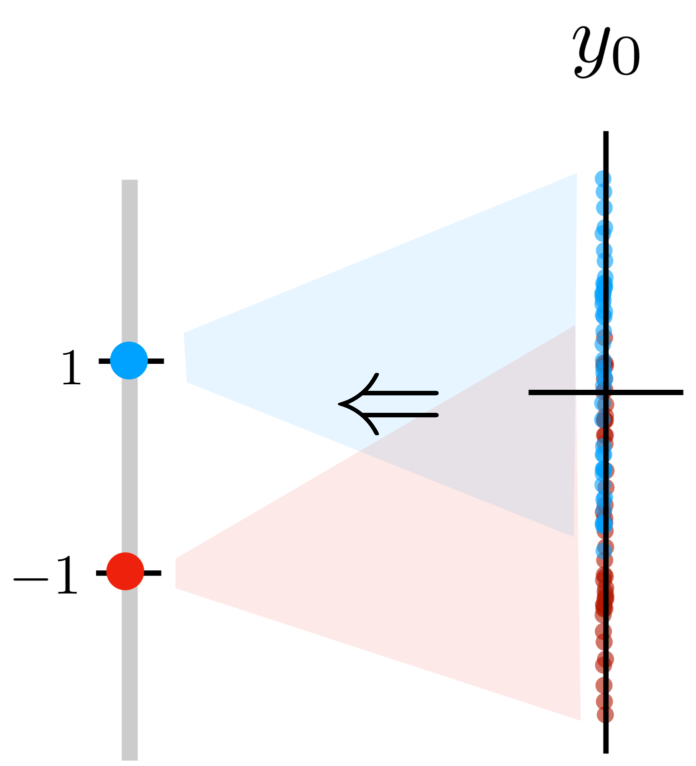
BASIS FUNCTIONS

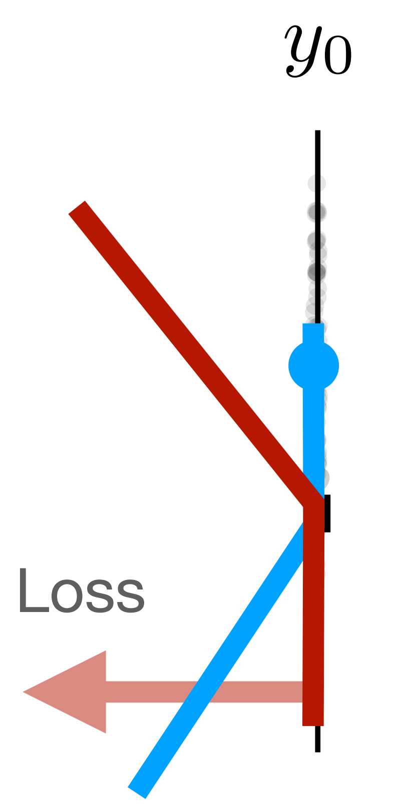
Hinge Loss

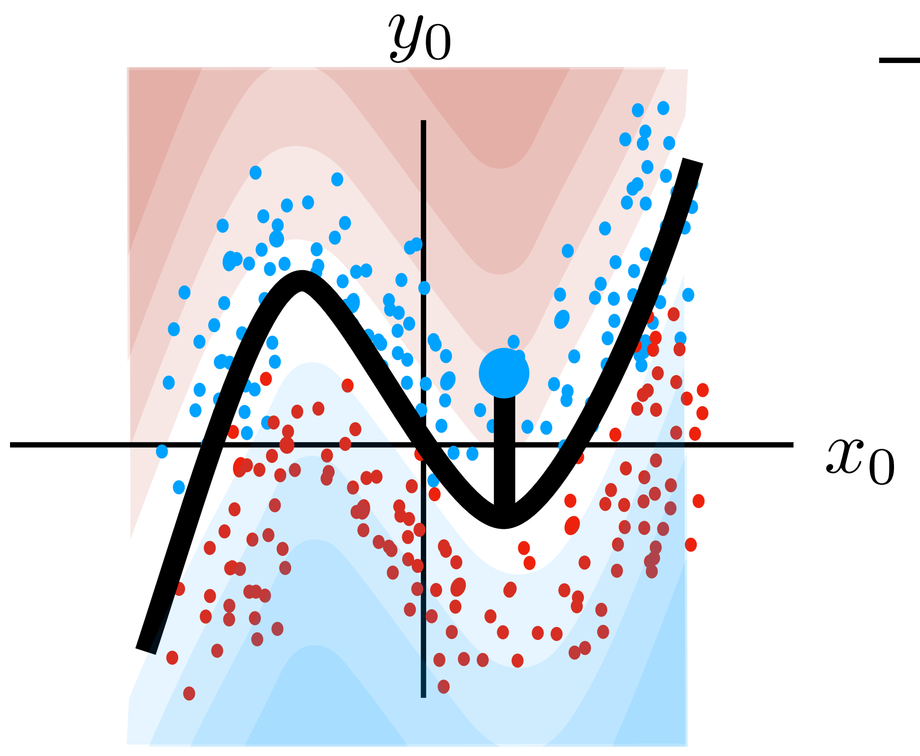
$$\sum_t \max\left\{0, \gamma_t\left(y_t - f(h_t(x_t, \xi_t))\right)\right\}$$

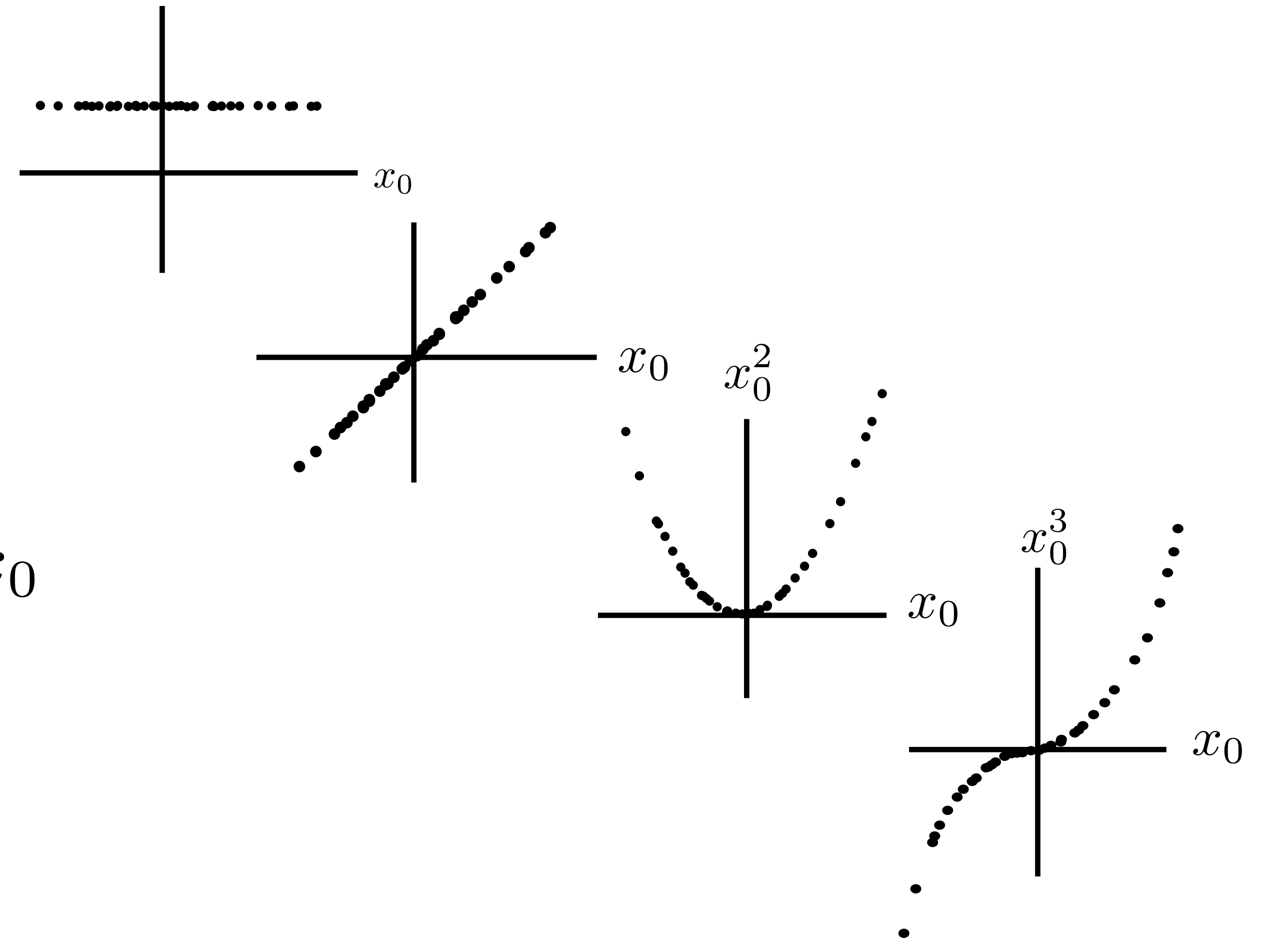
$$h_t(x_t) = [1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3]$$

$$\max\{0, (\cdot)\}$$

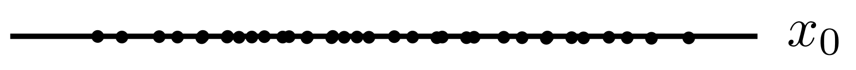








$\Leftarrow$



Loss Functions

OUTPUTS (Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS (Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Longleftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

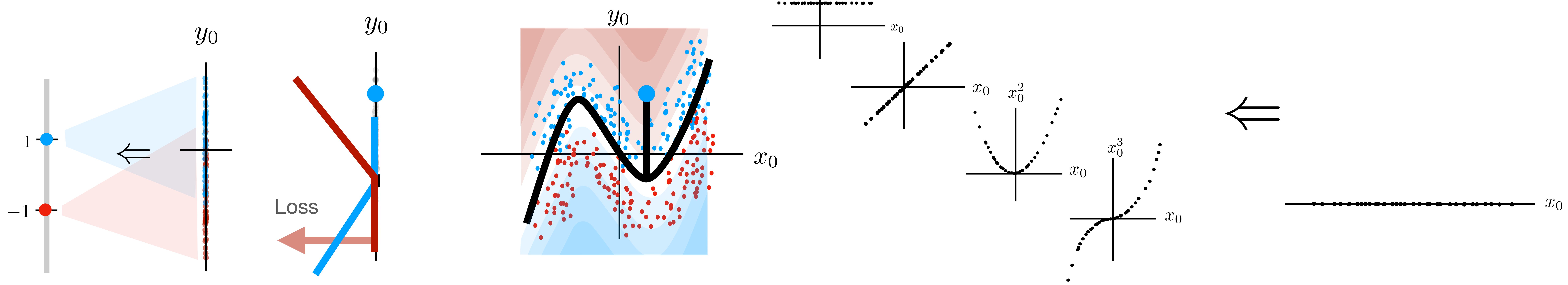
BASIS FUNCTIONS

Hinge Loss

$$\sum_t \max\left\{0, \gamma_t\left(y_t - f(h_t(x_t, \xi_t))\right)\right\}$$

$$h_t(x_t) = [1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3]$$

$\max\{0, (\cdot)\}$



Loss Functions

OUTPUTS
(Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0, \xi_0) & \cdots & h_{0n}(x_0, \xi_0) \\ h_{10}(x_1, \xi_1) & \cdots & h_{1n}(x_1, \xi_1) \\ h_{20}(x_2, \xi_2) & \cdots & h_{2n}(x_2, \xi_2) \\ h_{30}(x_3, \xi_3) & \cdots & h_{3n}(x_3, \xi_3) \\ h_{40}(x_4, \xi_4) & \cdots & h_{4n}(x_4, \xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T, \xi_T) & \cdots & h_{Tn}(x_T, \xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

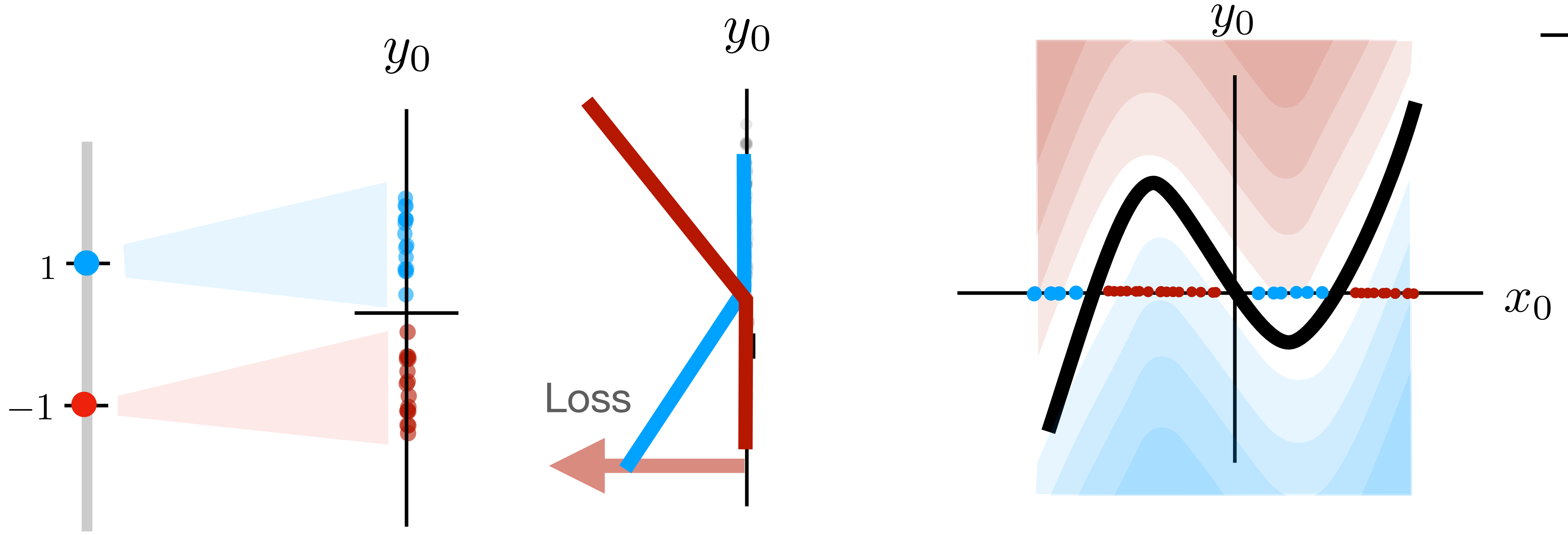
BASIS FUNCTIONS

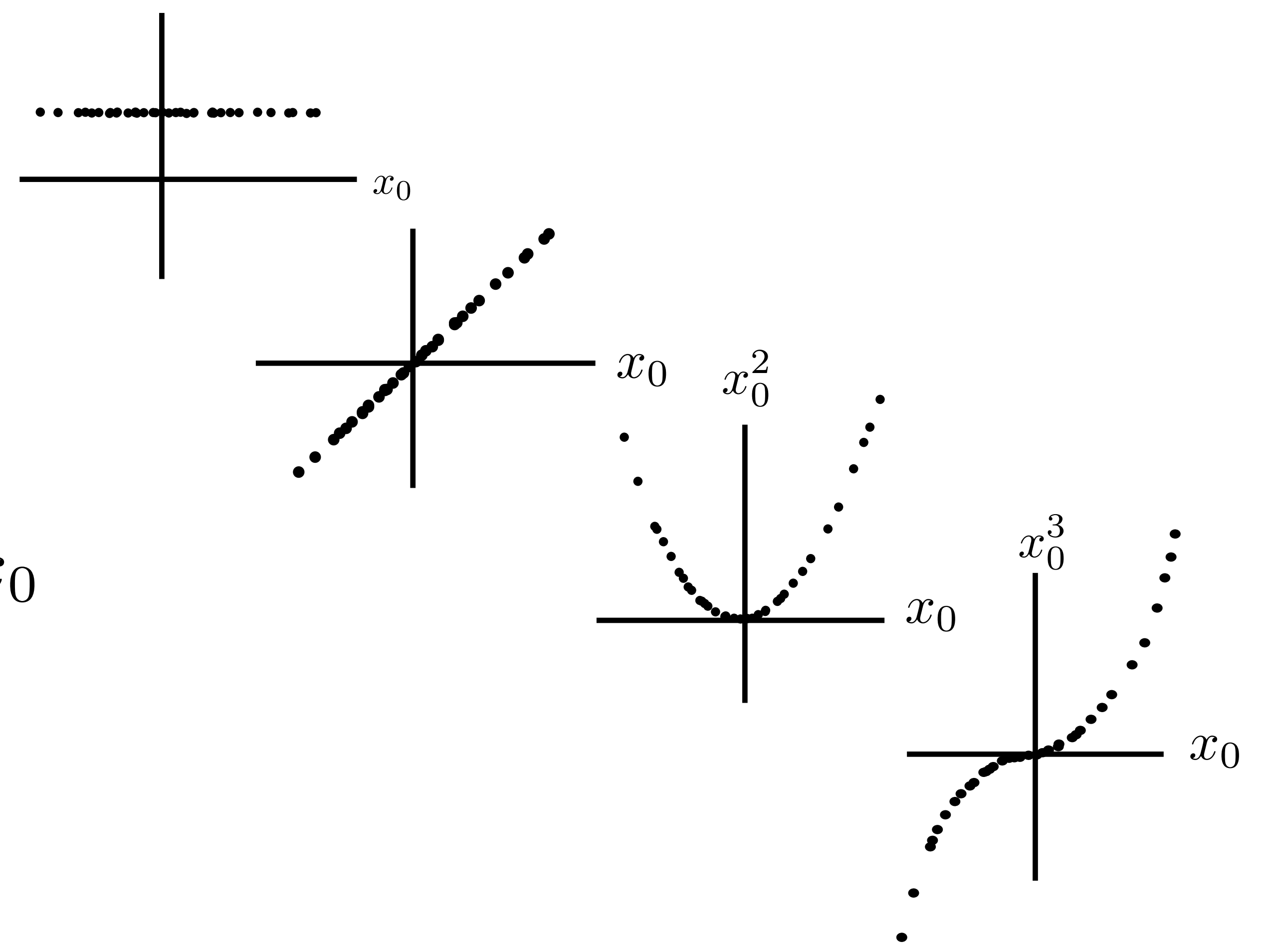
Hinge Loss

$$\sum_t \max\{0, \gamma_t f(h_t(x_t, \xi_t))\}$$

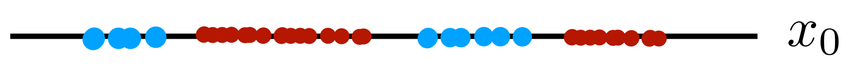
$$h_t(x_t) = [1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3]$$

$$\max\{0, (\cdot)\}$$





$\Leftarrow$



Loss Functions

OUTPUTS (Dependent Variables)

$$y_t = f(h_t(x_t, \xi_t))$$

BASIS FUNCTIONS

INPUTS (Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} h_{00}(x_0,\xi_0) & \cdots & h_{0n}(x_0,\xi_0) \\ h_{10}(x_1,\xi_1) & \cdots & h_{1n}(x_1,\xi_1) \\ h_{20}(x_2,\xi_2) & \cdots & h_{2n}(x_2,\xi_2) \\ h_{30}(x_3,\xi_3) & \cdots & h_{3n}(x_3,\xi_3) \\ h_{40}(x_4,\xi_4) & \cdots & h_{4n}(x_4,\xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T,\xi_T) & \cdots & h_{Tn}(x_T,\xi_T) \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} & \xi_{00} & \cdots & \xi_{0n'} \\ x_{10} & \cdots & x_{1n} & \xi_{10} & \cdots & \xi_{1n'} \\ x_{20} & \cdots & x_{2n} & \xi_{20} & \cdots & \xi_{2n'} \\ x_{30} & \cdots & x_{3n} & \xi_{30} & \cdots & \xi_{3n'} \\ x_{40} & \cdots & x_{4n} & \xi_{40} & \cdots & \xi_{4n'} \\ \vdots & & \vdots & \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} & \xi_{T0} & \cdots & \xi_{Tn'} \end{bmatrix}$$

LOSS FUNCTIONS

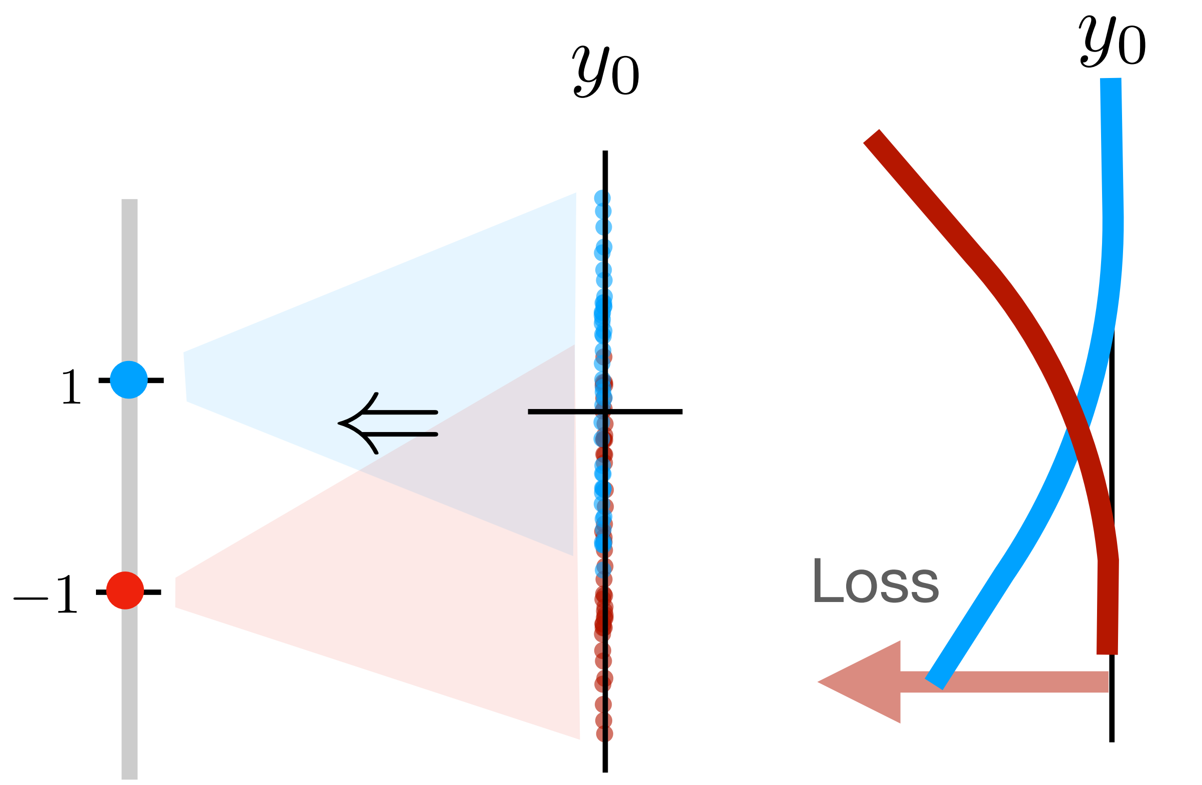
BASIS FUNCTIONS

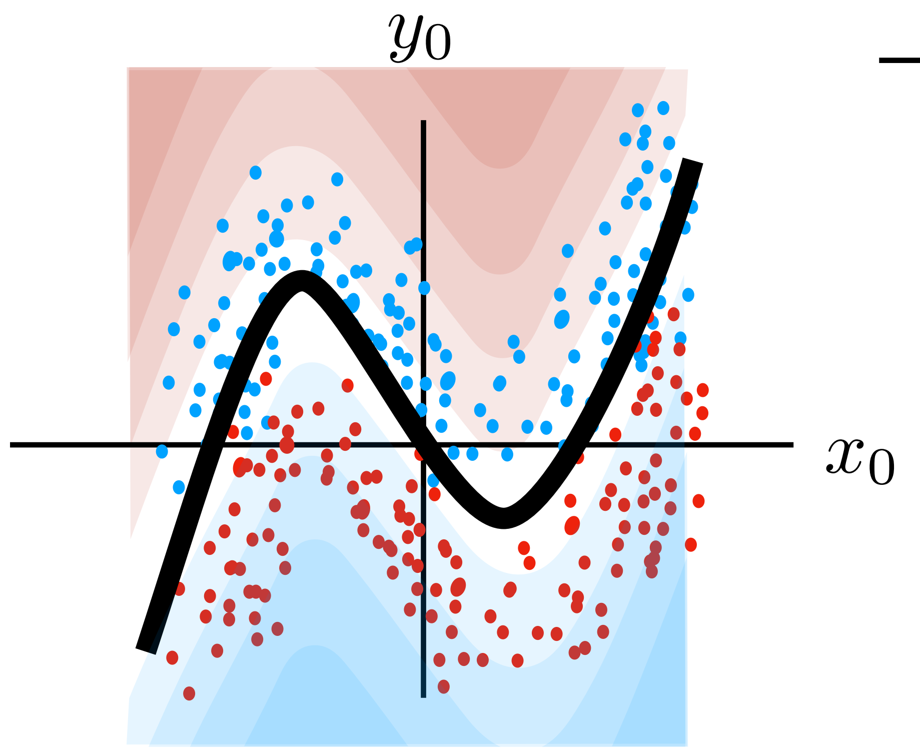
Smoothed Hinge Loss

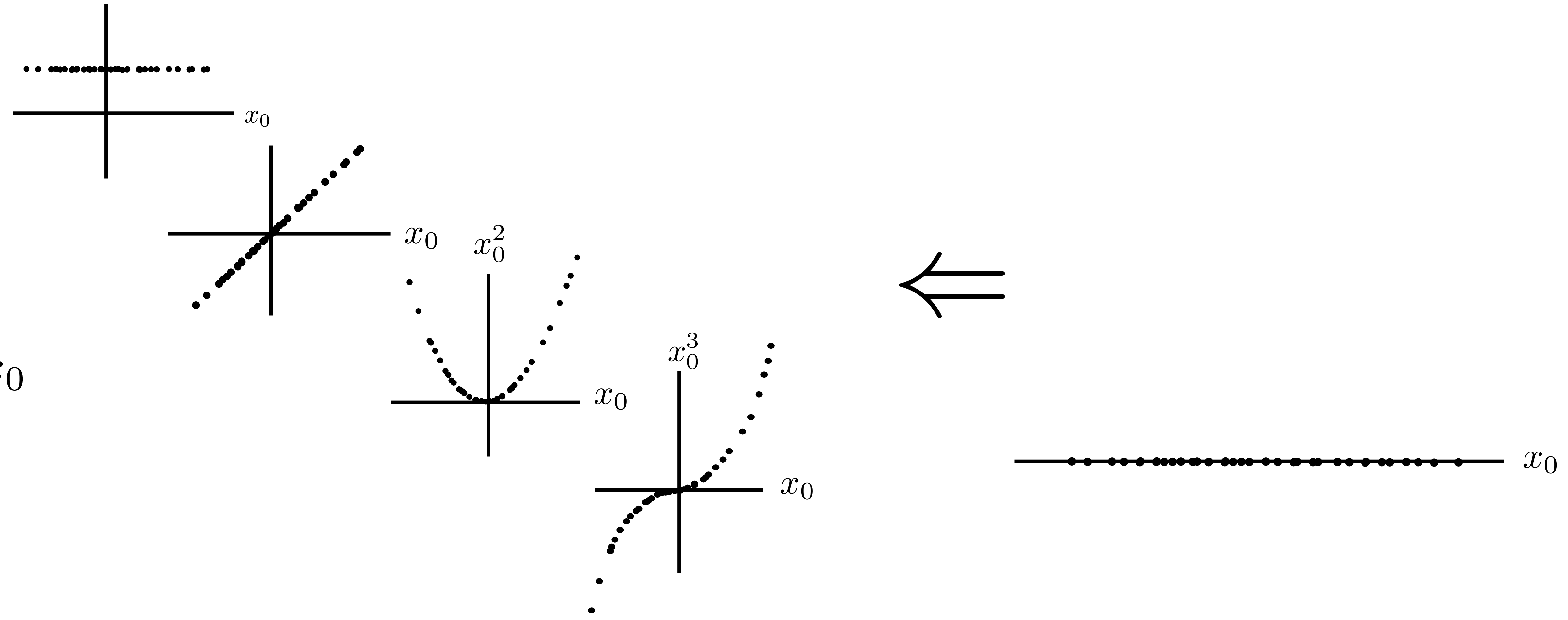
$$\ln\left(e^{(\cdot)} + 1\right)$$

$$\sum_t \ln\left(e^{\gamma_t\left(y_t - f(h_t(x_t,\xi_t))\right)} + 1\right)$$

$$h_t(x_t) = [1 \ x_{t0} \ x_{t0}^2 \ x_{t0}^3]$$







Linear Regression

OUTPUTS
(Dependent Variables)

$$y_t = \theta^T x_t$$

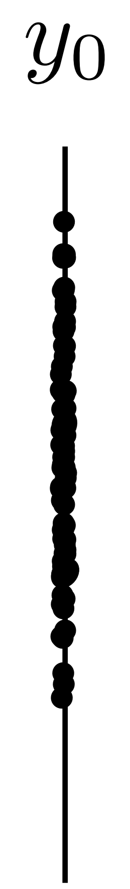
INPUTS
(Independent Variables)

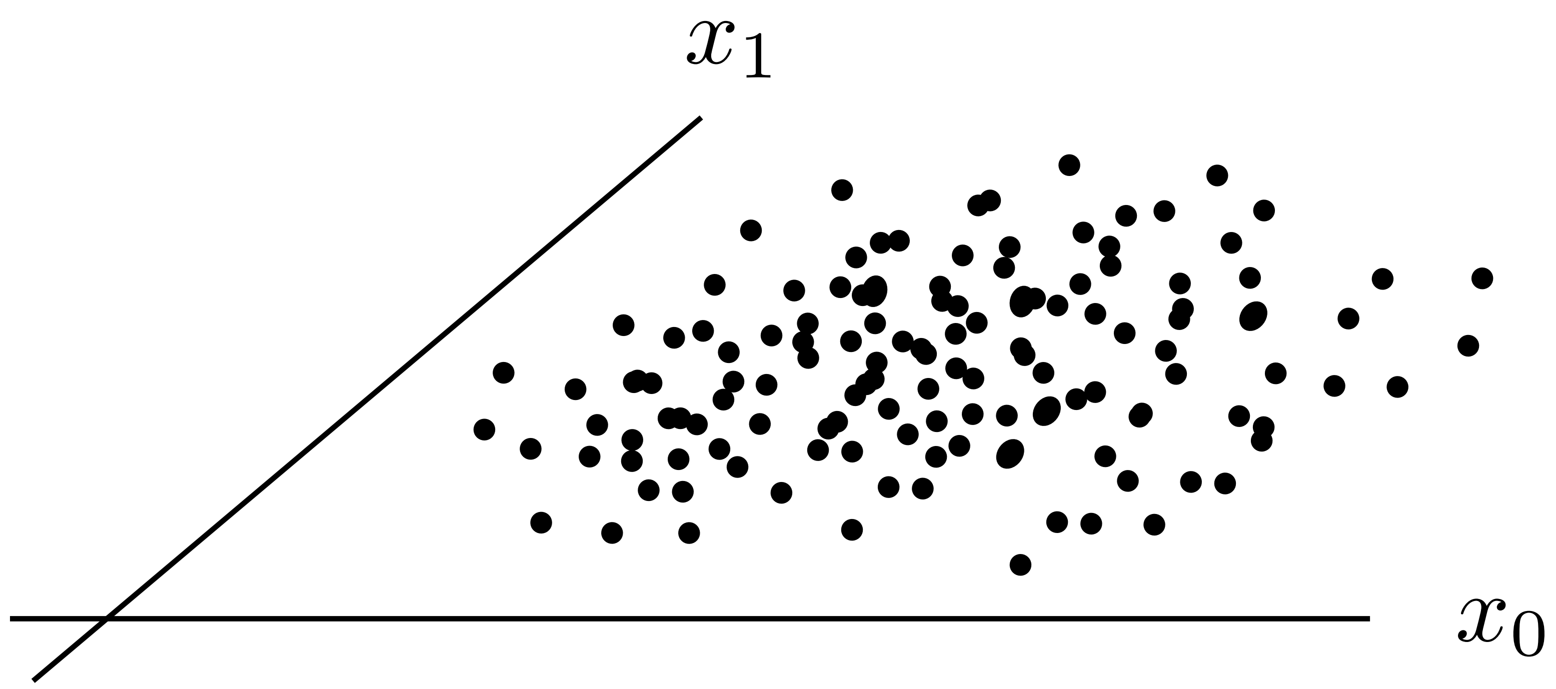
$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_n \end{bmatrix}$$

$$y = X\theta$$

COST: $$\min \left|\left|y - X\theta\right|\right|^2$$

SOLN: $$\theta = (X^T X)^{-1} X^T y$$





Linear Regression

OUTPUTS
(Dependent Variables)

$$y_t = \theta^T x_t$$

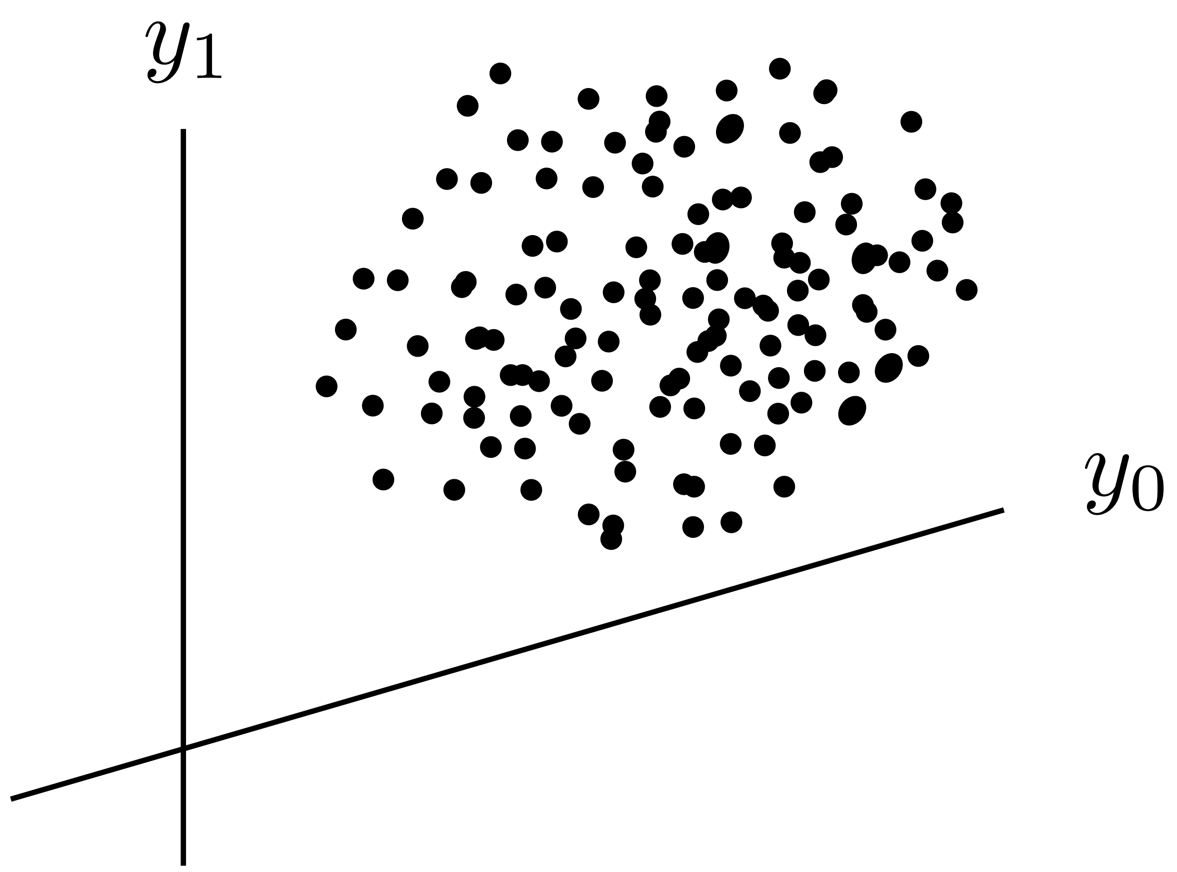
INPUTS
(Independent Variables)

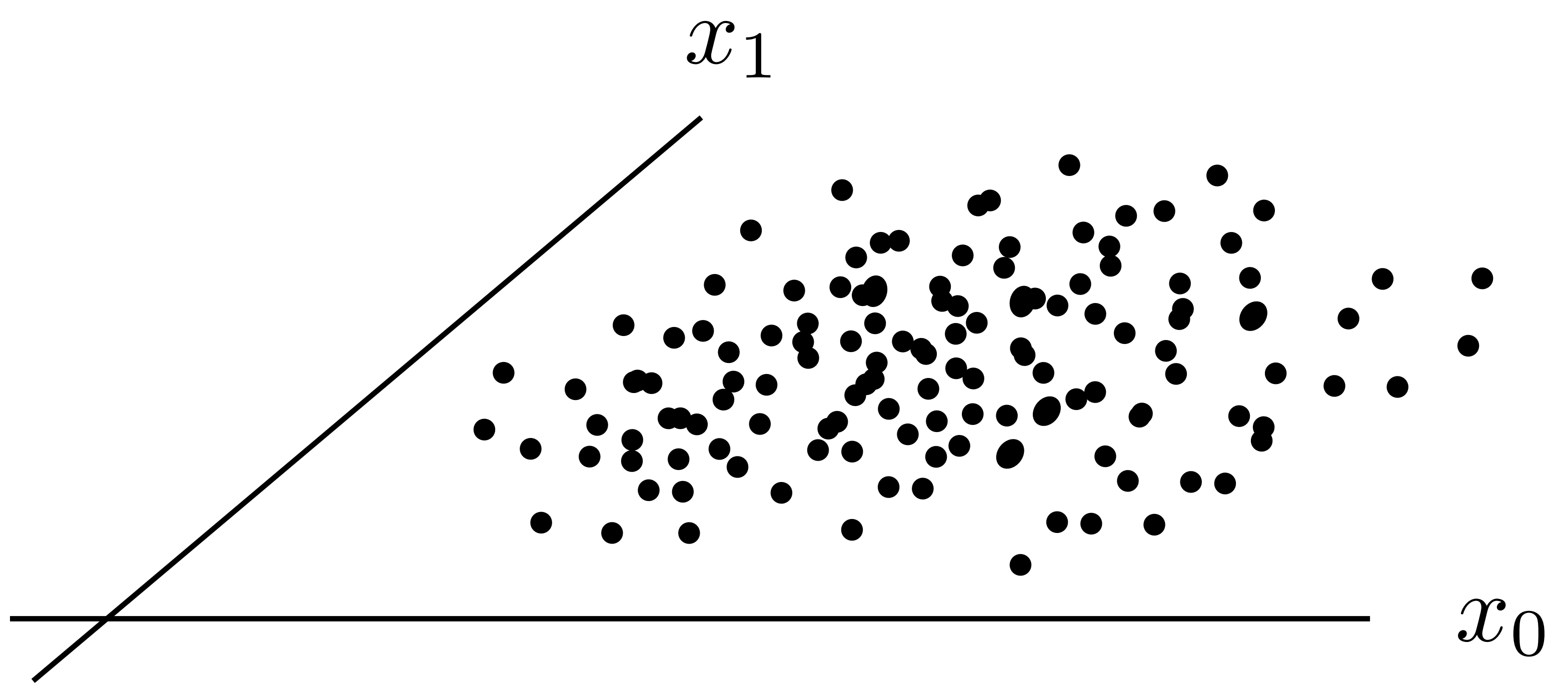
$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix} \begin{bmatrix} \theta_{00} & \cdots & \theta_{0m} \\ \vdots & & \vdots \\ \theta_{n0} & \cdots & \theta_{nm} \end{bmatrix}$$

$$y = X\theta$$

COST: $\min \; \left|\left|y - X\theta\right|\right|^2$

SOLN: $\theta = (X^T X)^{-1} X^T y$





Linear Regression

OUTPUTS
(Dependent Variables)

$$y_t = \theta_t^T h_t(x_t)$$

BASIS FUNCTIONS

INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} h_{00}(x_0,\xi_0) & \cdots & h_{0n}(x_0,\xi_0) \\ h_{10}(x_1,\xi_1) & \cdots & h_{1n}(x_1,\xi_1) \\ h_{20}(x_2,\xi_2) & \cdots & h_{2n}(x_2,\xi_2) \\ h_{30}(x_3,\xi_3) & \cdots & h_{3n}(x_3,\xi_3) \\ h_{40}(x_4,\xi_4) & \cdots & h_{4n}(x_4,\xi_4) \\ \vdots & & \vdots \\ h_{T0}(x_T,\xi_T) & \cdots & h_{Tn}(x_T,\xi_T) \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_n \end{bmatrix} \Longleftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix}$$

$$y = h(X)\theta$$

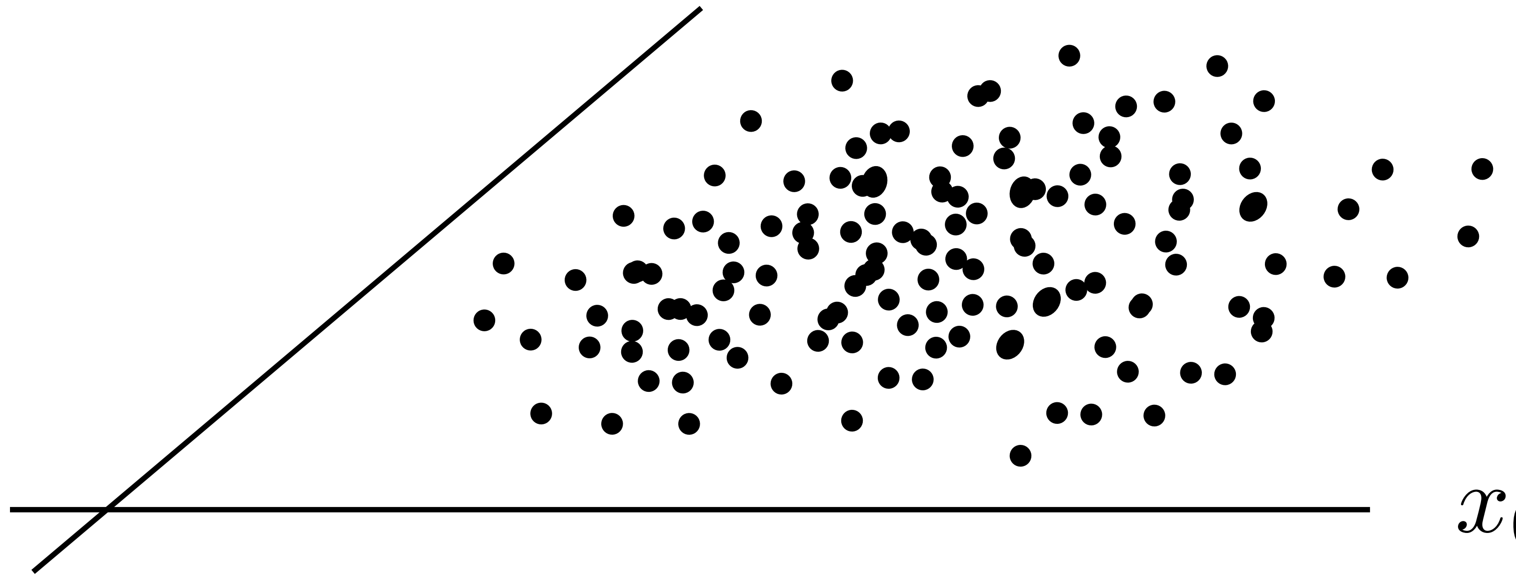
COST: $$\min \; \big|\big|y - h(X)\theta\big|\big|^2$$

BASIS FUNCTIONS

SOLN: $$\theta = \big(h(X)^T h(X)\big)^{-1} h(X)^T y$$

y_0

x_1



x_0

Linear Regression

OUTPUTS (Dependent Variables)

$$y_t = \theta_t^T h_t(x_t)$$

BASIS FUNCTIONS

INPUTS (Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} 1 & x_{00} & x_{00}^2 & x_{00}^3 \\ 1 & x_{10} & x_{10}^2 & x_{10}^3 \\ 1 & x_{20} & x_{20}^2 & x_{20}^3 \\ 1 & x_{30} & x_{30}^2 & x_{30}^3 \\ \vdots & & \vdots & \\ 1 & x_{T0} & x_{T0}^2 & x_{T0}^3 \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_n \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix}$$

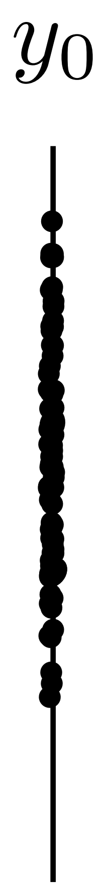
$$y = h(X)\theta$$

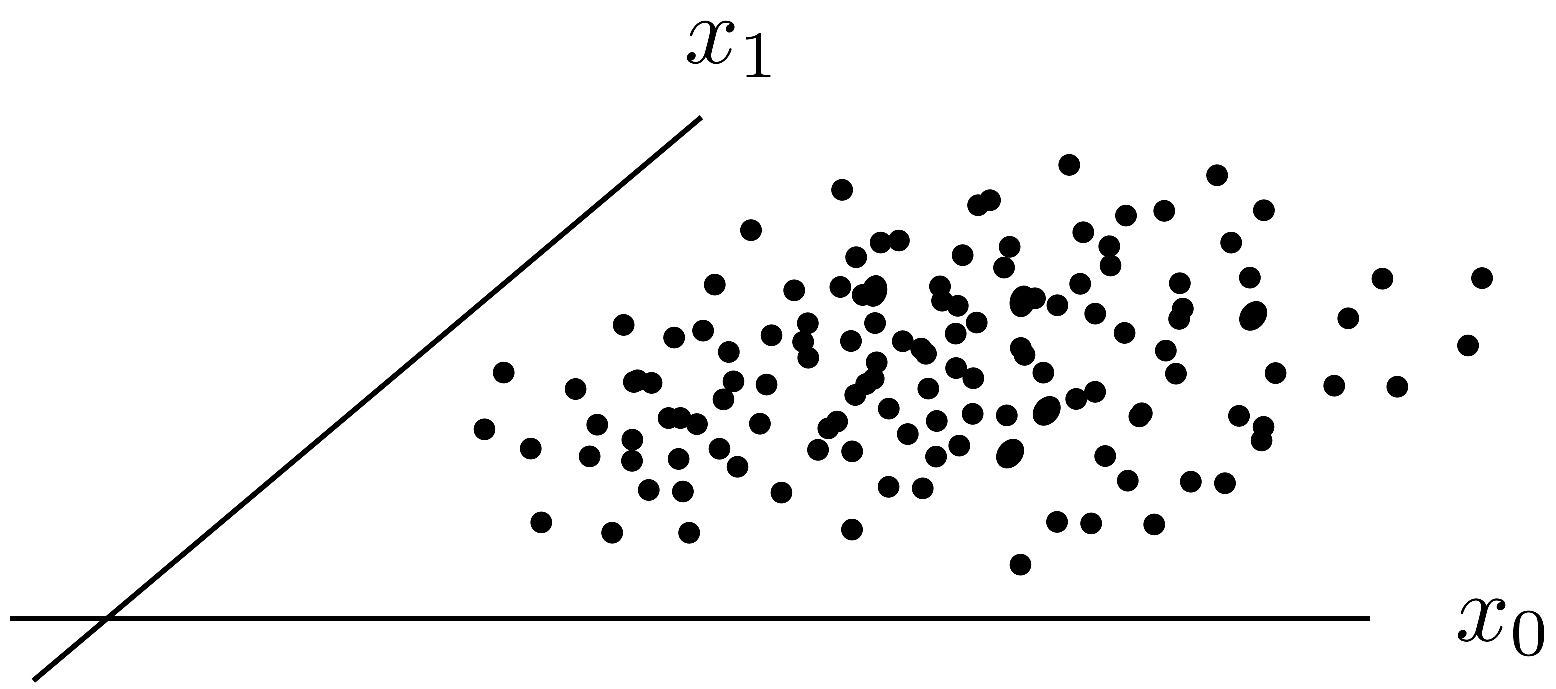
COST: $\min \; \left|\left| y - h(X)\theta \right|\right|^2$

BASIS FUNCTIONS

$$h_t(x_t) = \left[1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3\right]$$

SOLN: $\theta = \left(h(X)^T h(X)\right)^{-1} h(X)^T y$





Linear Regression

OUTPUTS
(Dependent Variables)

$$y_t = \theta_t^T h_t(x_t)$$

BASIS FUNCTIONS

INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} 1 & x_{00} & x_{00}^2 & x_{00}^3 \\ 1 & x_{10} & x_{10}^2 & x_{10}^3 \\ 1 & x_{20} & x_{20}^2 & x_{20}^3 \\ 1 & x_{30} & x_{30}^2 & x_{30}^3 \\ \vdots & & \vdots & \\ 1 & x_{T0} & x_{T0}^2 & x_{T0}^3 \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_n \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix}$$

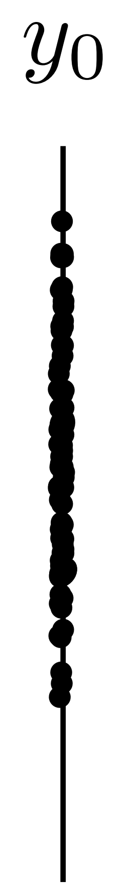
$$y = h(X)\theta$$

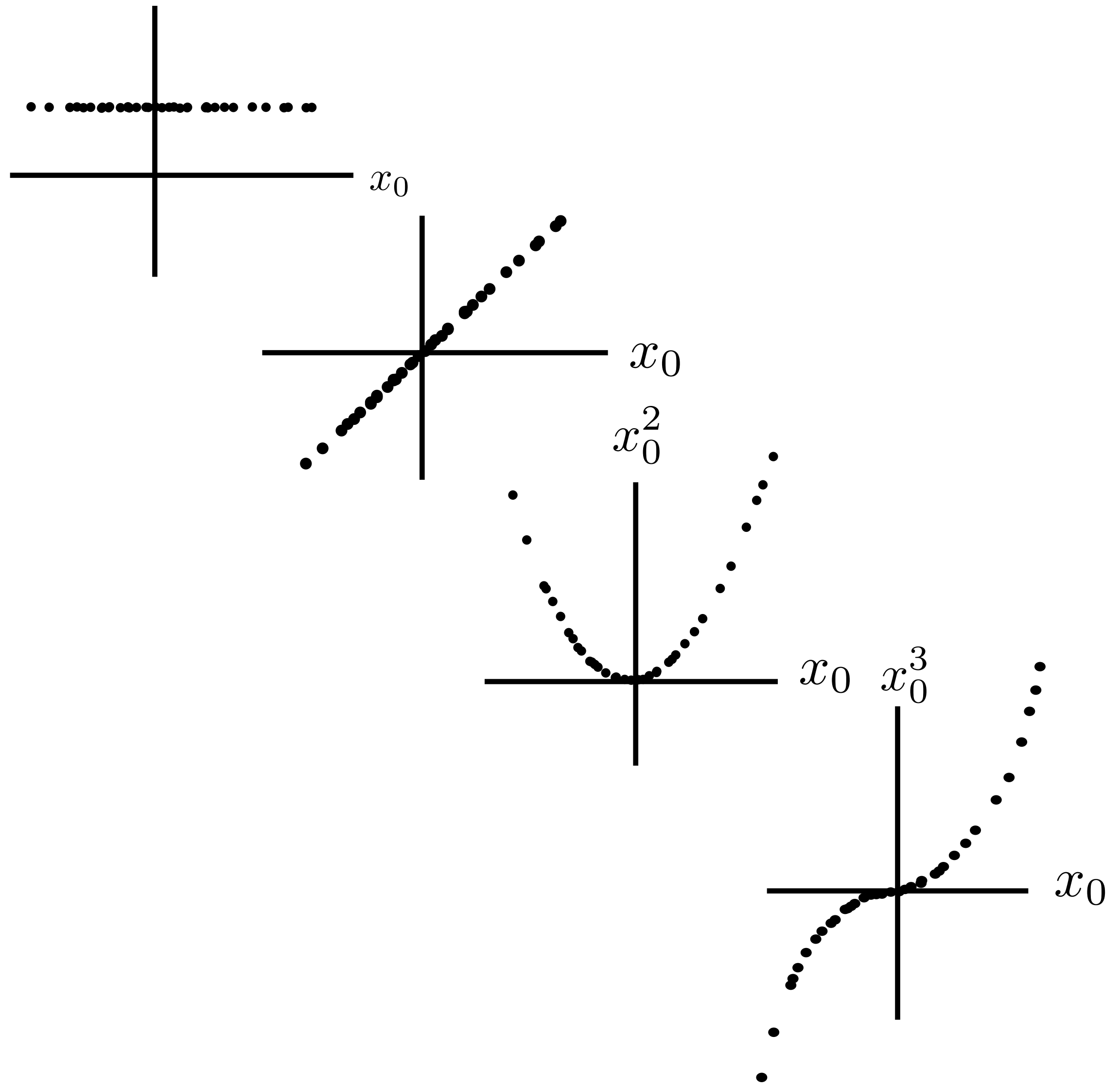
COST: $$\min \; \left|\left| y - h(X)\theta \right|\right|^2$$

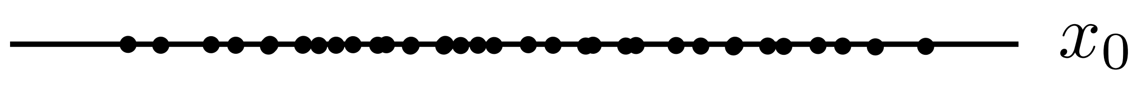
BASIS FUNCTIONS

$$h_t(x_t) = \left[1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3\right]$$

SOLN: $$\theta = \left(h(X)^T h(X)\right)^{-1} h(X)^T y$$







Linear Regression

OUTPUTS (Dependent Variables)

$$y_t = \theta_t^T h_t(x_t)$$

BASIS FUNCTIONS

INPUTS (Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} 1 & x_{00} & x_{00}^2 & x_{00}^3 \\ 1 & x_{10} & x_{10}^2 & x_{10}^3 \\ 1 & x_{20} & x_{20}^2 & x_{20}^3 \\ 1 & x_{30} & x_{30}^2 & x_{30}^3 \\ \vdots & & \vdots & \\ 1 & x_{T0} & x_{T0}^2 & x_{T0}^3 \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_n \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix}$$

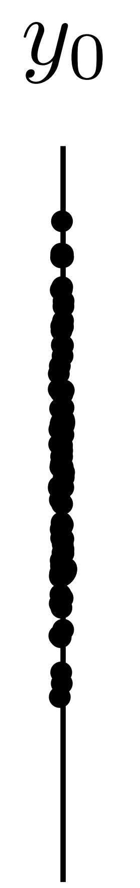
$$y = h(X)\theta$$

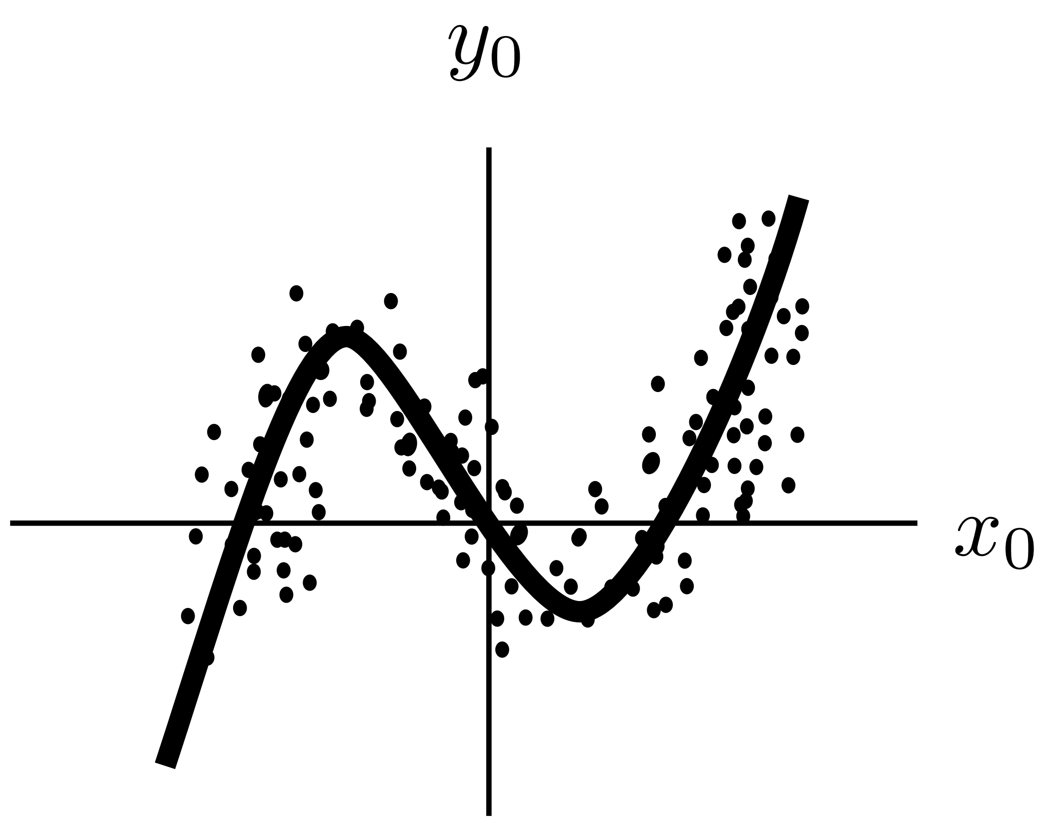
COST: $\min \left|\left| y - h(X)\theta \right|\right|^2$

BASIS FUNCTIONS

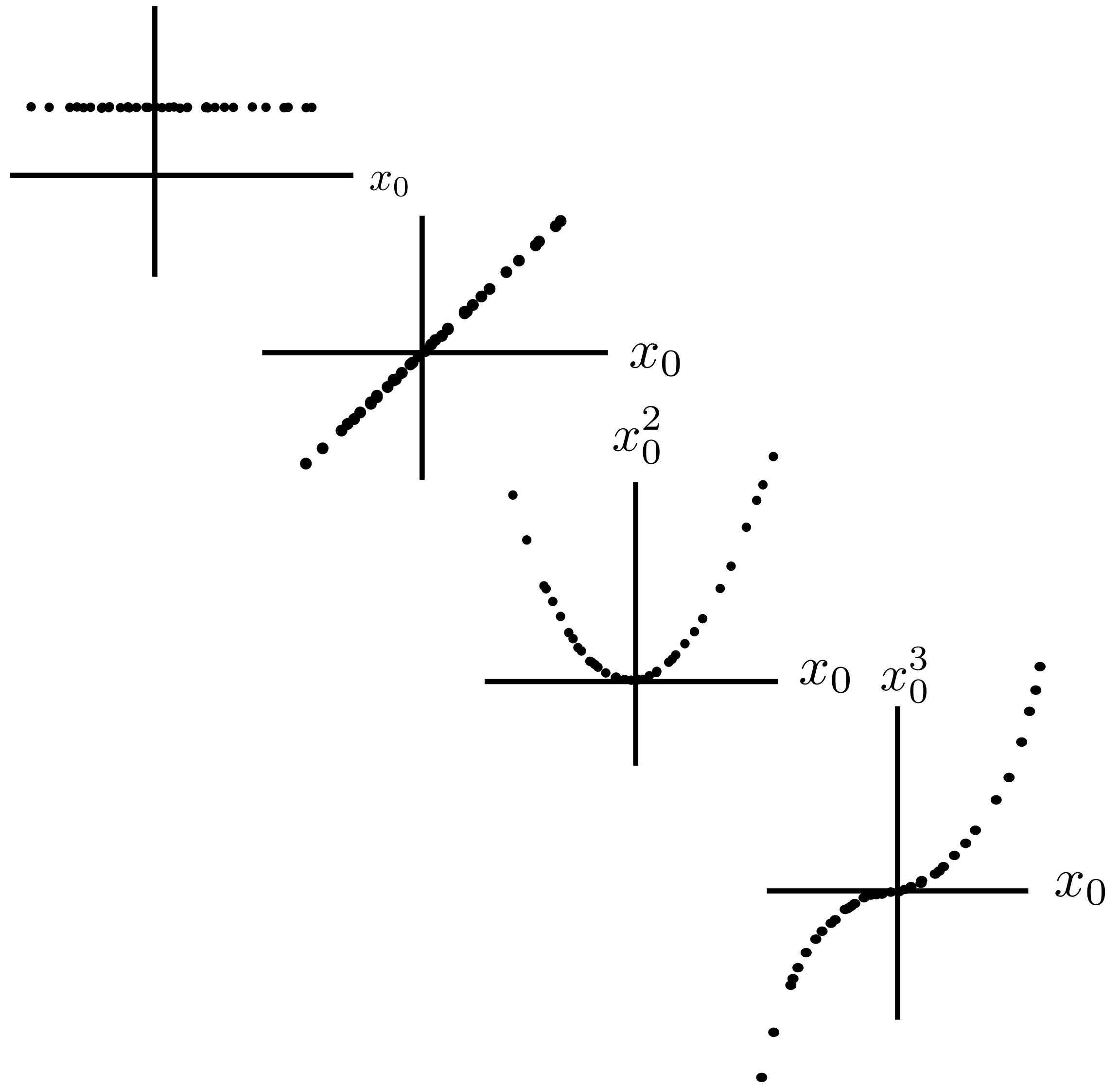
$$h_t(x_t) = \left[1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3\right]$$

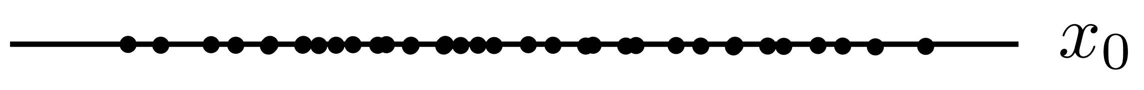
SOLN: $\theta = \left(h(X)^T h(X)\right)^{-1} h(X)^T y$





$\Leftarrow$





Linear Regression

OUTPUTS
(Dependent Variables)

$$y_t = \theta_t^T h_t(x_t)$$

BASIS FUNCTIONS

INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} 1 & x_{00} & x_{01} & x_{00}^2 & x_{00}x_{01} & x_{01}^2 \\ 1 & x_{10} & x_{11} & x_{10}^2 & x_{10}x_{11} & x_{11}^2 \\ 1 & x_{20} & x_{21} & x_{20}^2 & x_{20}x_{21} & x_{21}^2 \\ 1 & x_{30} & x_{31} & x_{30}^2 & x_{30}x_{31} & x_{31}^2 \\ \vdots & & \vdots & & & \\ 1 & x_{T0} & x_{T1} & x_{T0}^2 & x_{T0}x_{T1} & x_{T1}^2 \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_n \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix}$$

$$y = h(X)\theta$$

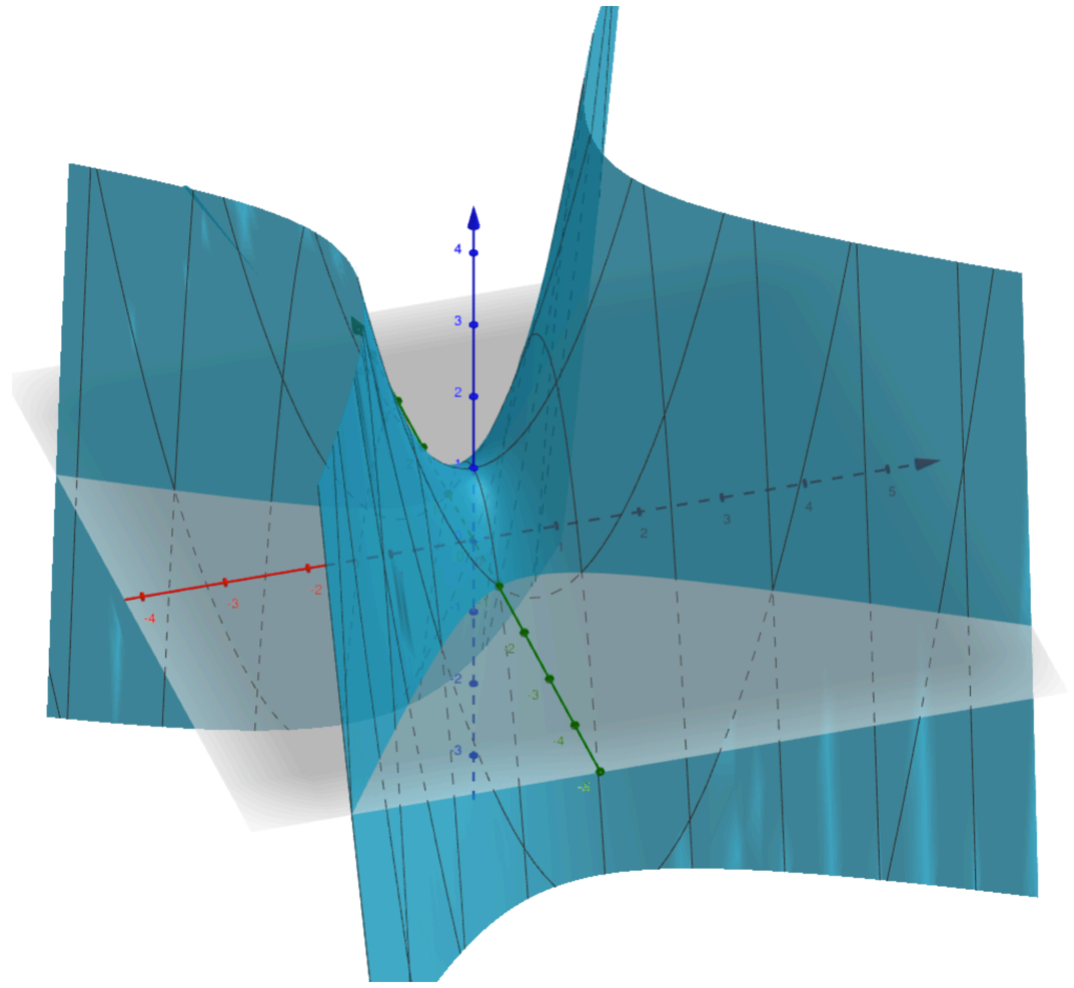
COST: $\min \; ||y - h(X)\theta||^2$

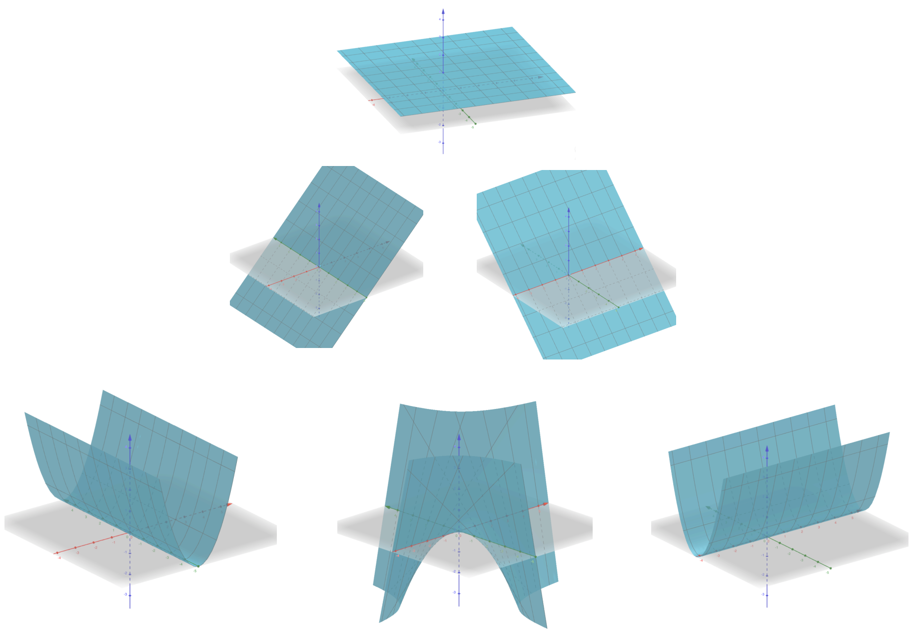
BASIS FUNCTIONS

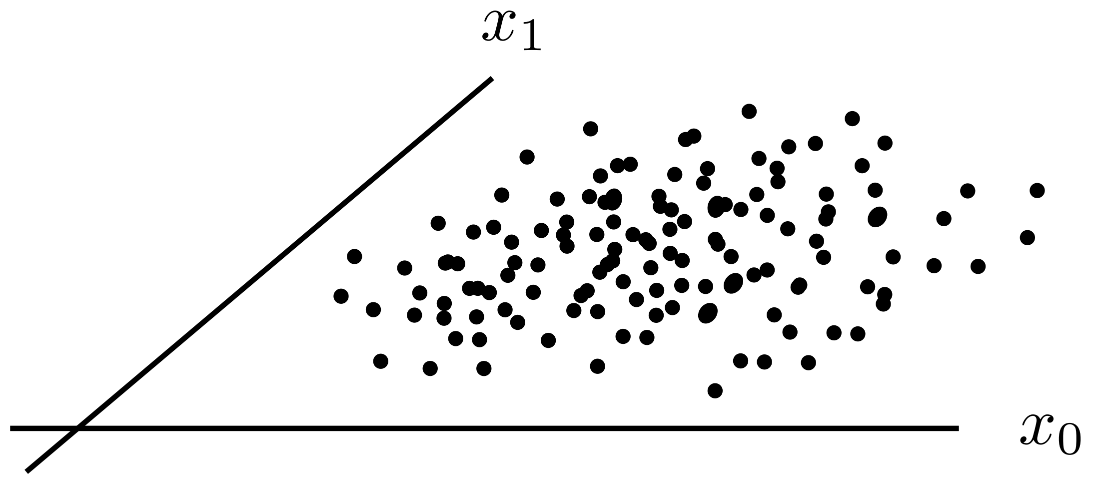
$$h_t(x_t) = [1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3]$$

SOLN: $\theta = \left(h(X)^T h(X)\right)^{-1} h(X)^T y$

y_0







Linear Regression

OUTPUTS (Dependent Variables)

$$y_t = \theta_t^T h_t(x_t)$$

BASIS FUNCTIONS

INPUTS (Independent Variables)

$$\begin{bmatrix} y_{00} & \cdots & y_{0m} \\ y_{10} & \cdots & y_{1m} \\ y_{20} & \cdots & y_{2m} \\ y_{30} & \cdots & y_{3m} \\ y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots \\ y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \quad y = h(X)\theta$$

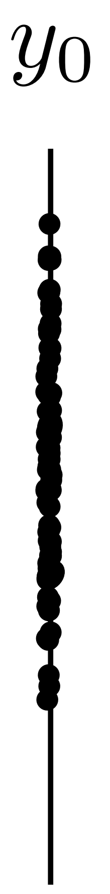
$$\begin{bmatrix} 1 & x_{00} & x_{01} & x_{00}^2 & x_{00}x_{01} & x_{01}^2 & x_{00}^3 & x_{00}^2x_{01} & x_{01}^2x_{00} & x_{01}^3 \\ 1 & x_{10} & x_{11} & x_{10}^2 & x_{10}x_{11} & x_{11}^2 & x_{10}^3 & x_{10}^2x_{11} & x_{11}^2x_{10} & x_{11}^3 \\ 1 & x_{20} & x_{21} & x_{20}^2 & x_{20}x_{21} & x_{21}^2 & x_{20}^3 & x_{20}^2x_{21} & x_{21}^2x_{20} & x_{21}^3 \\ 1 & x_{30} & x_{31} & x_{30}^2 & x_{30}x_{31} & x_{31}^2 & x_{30}^3 & x_{30}^2x_{31} & x_{31}^2x_{30} & x_{31}^3 \\ \vdots & & \vdots \\ 1 & x_{T0} & x_{T1} & x_{T0}^2 & x_{T0}x_{T1} & x_{T1}^2 & x_{T0}^3 & x_{T0}^2x_{T1} & x_{T1}^2x_{T0} & x_{T1}^3 \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_n \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix}$$

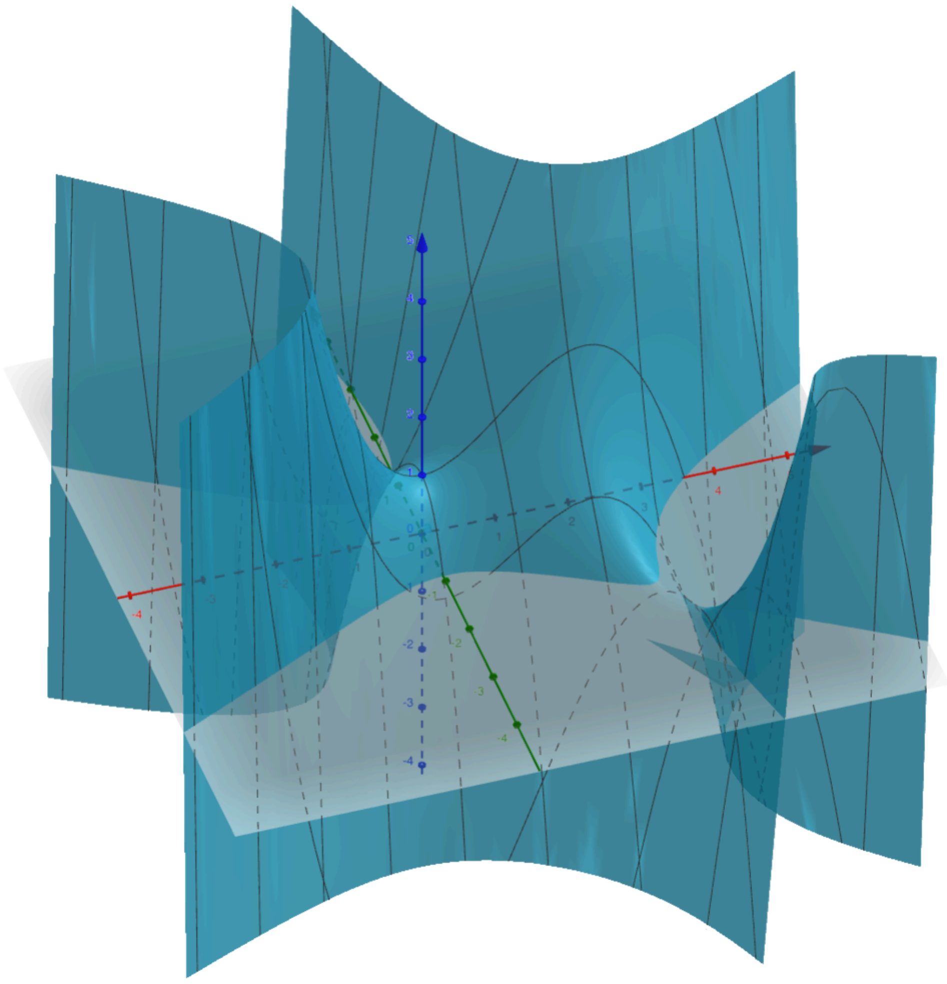
COST: $$\min \; \left|\left|y - h(X)\theta\right|\right|^2$$

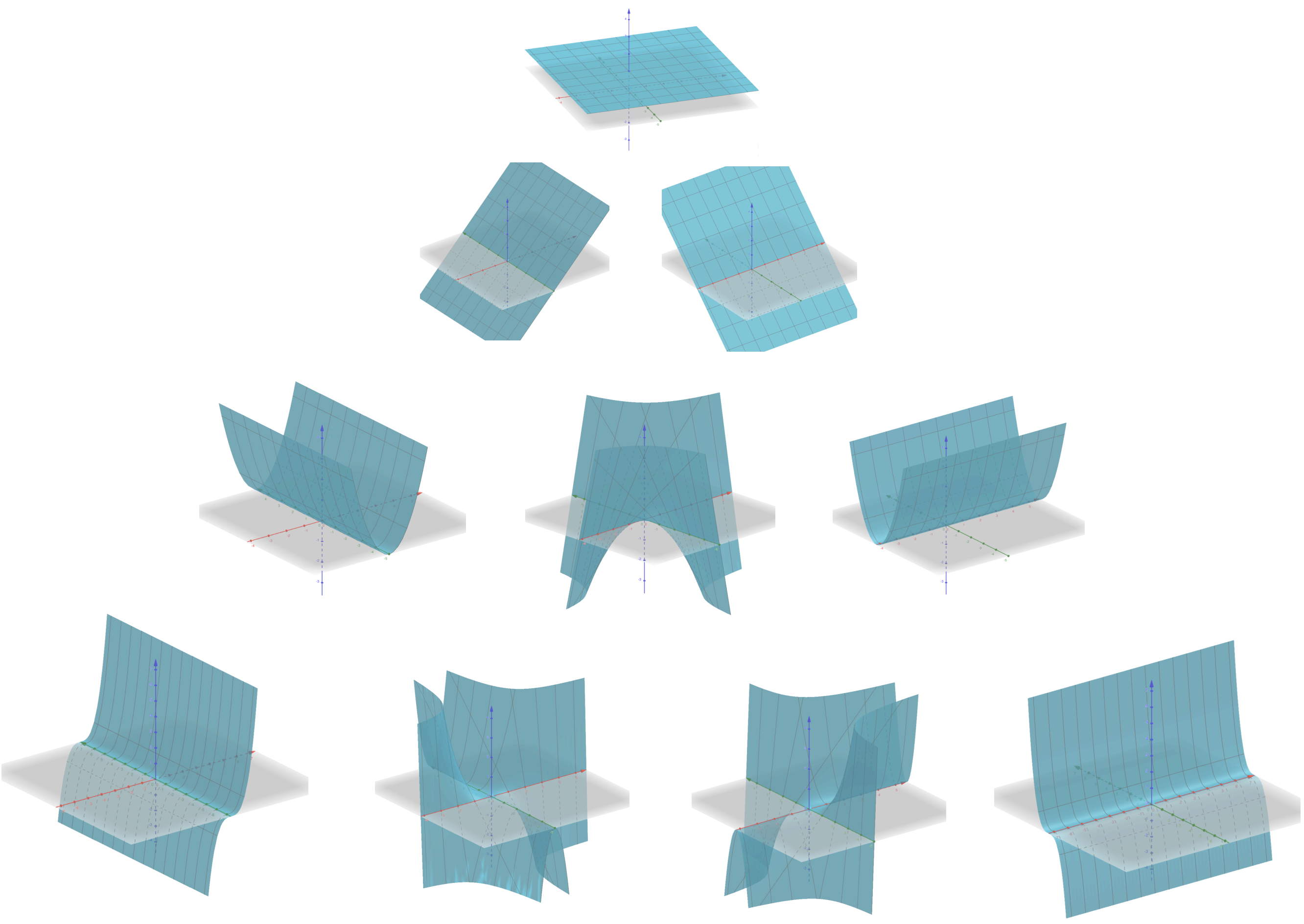
BASIS FUNCTIONS

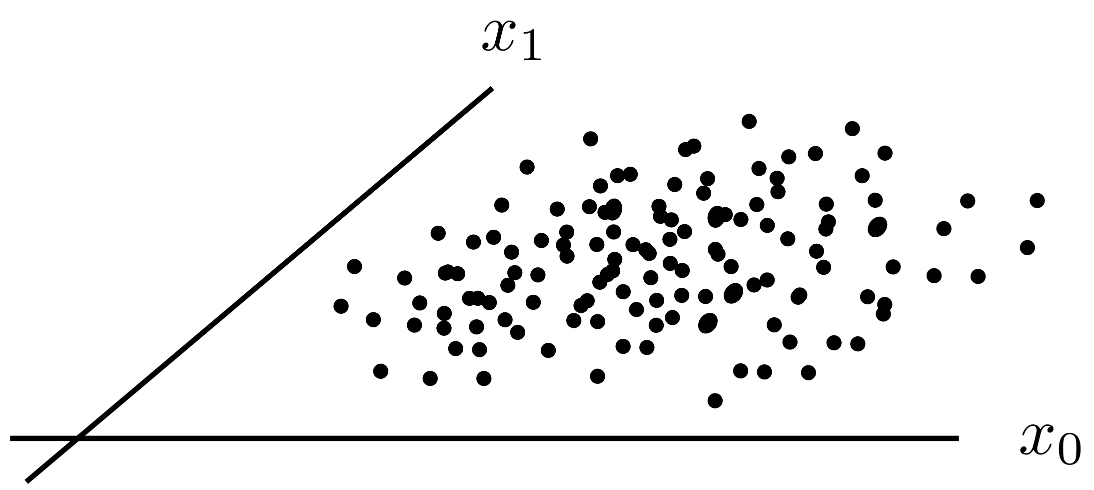
$$h_t(x_t) = \left[1 \; x_{t0} \; x_{t0}^2 \; x_{t0}^3\right]$$

SOLN: $$\theta = \left(h(X)^T h(X)\right)^{-1} h(X)^T y$$









Linear Regression

OUTPUTS (Dependent Variables)

$$y_t = \theta^T x_t + v_t$$

$$v_t \sim \mathcal{N}(0, \sigma_t^2)$$

INPUTS (Independent Variables)

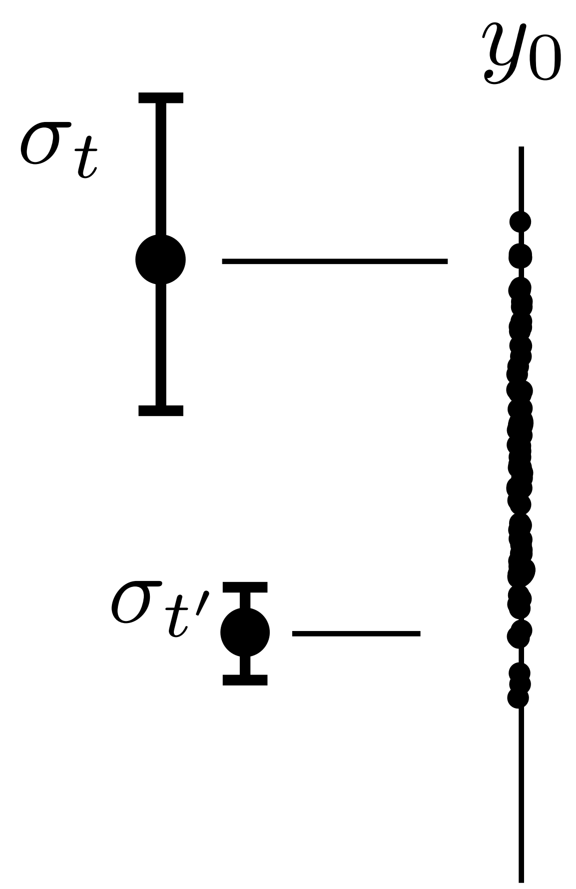
$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_n \end{bmatrix}$$

$$y = X\theta + v$$

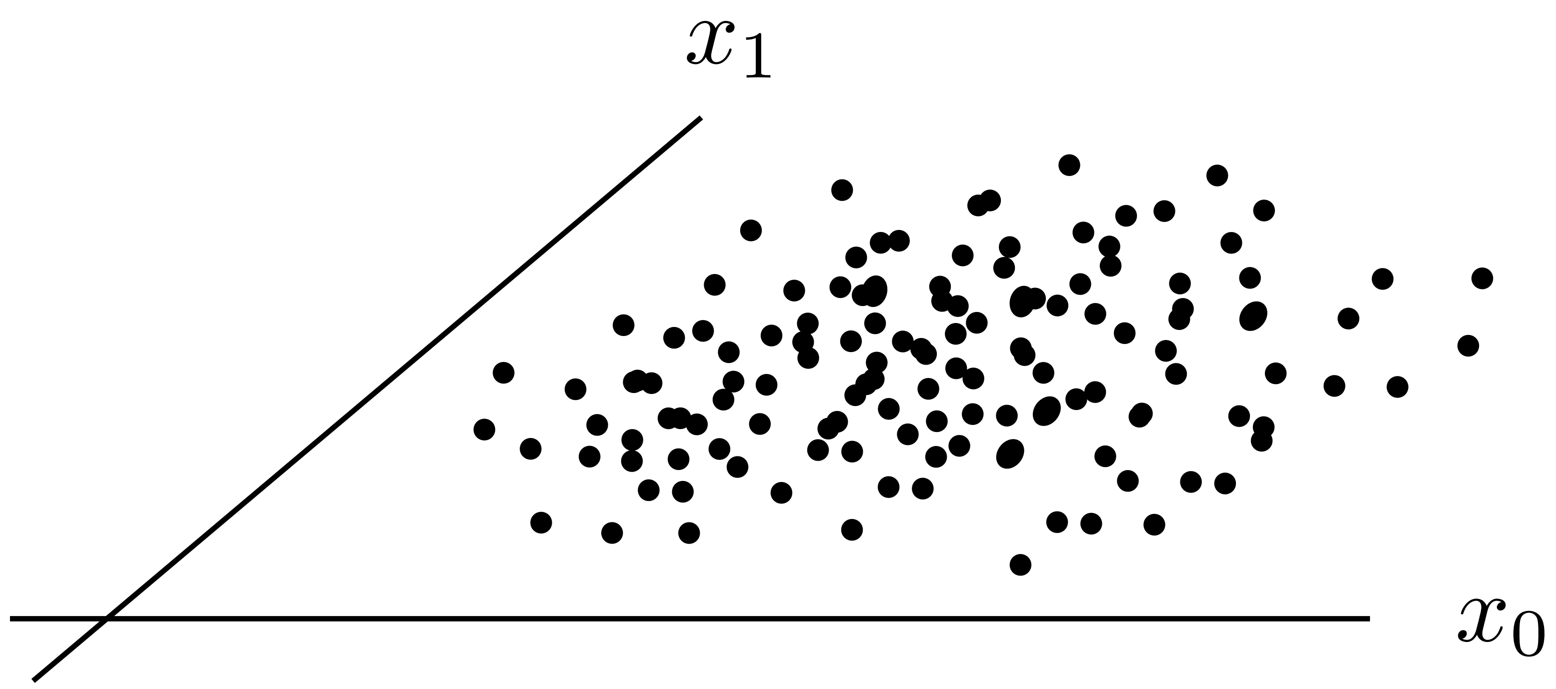
COST: $$\min \quad (y - X\theta)^T W^{-1} (y - X\theta)$$

WEIGHTS: $$W = \mathrm{diag}(\sigma_0^2, \ldots, \sigma_T^2)$$

SOLN: $$\theta = (X^T W^{-1} X)^{-1} X W^{-1} y$$



Weighted Least Squares



Linear Regression

OUTPUTS
(Dependent Variables)

$$y_t = \theta^T x_t + v_t$$

$$v_t \sim \mathcal{N}(0, \sigma_t^2)$$

INPUTS
(Independent Variables)

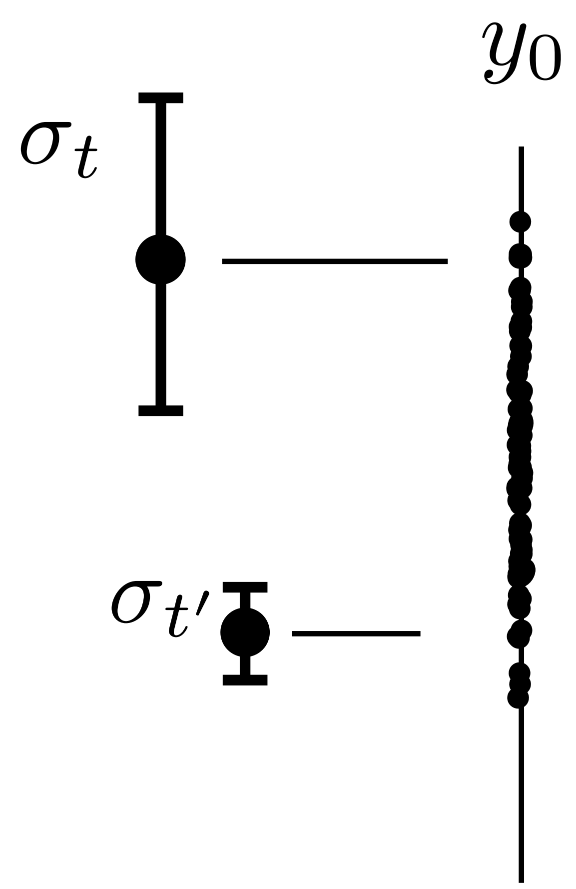
$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_n \end{bmatrix}$$

$$y = X\theta + v$$

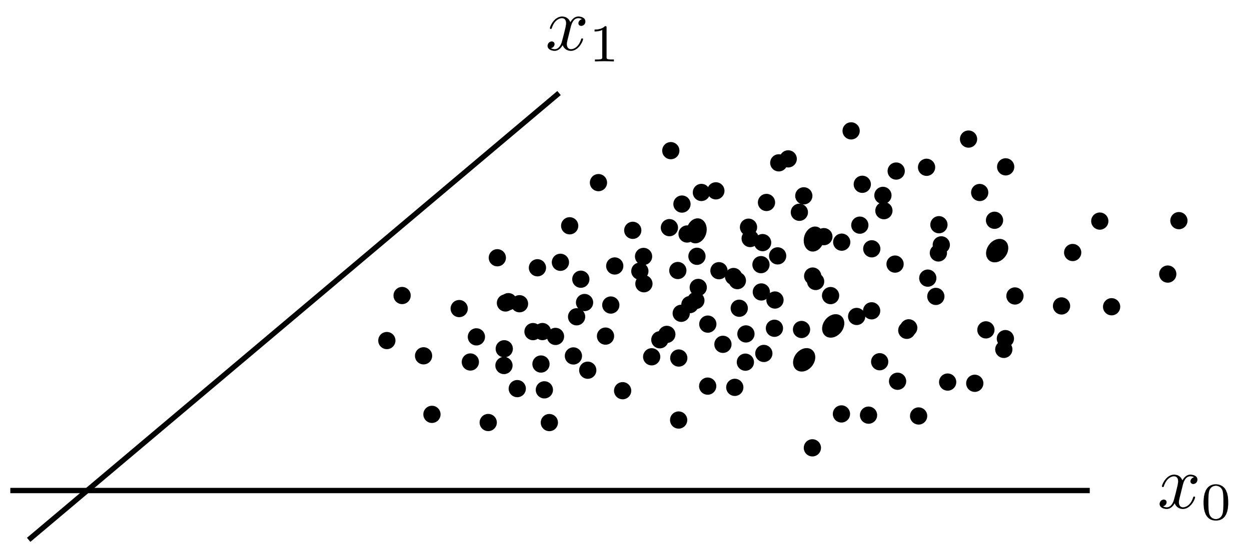
COST: $$\min \quad (y - X\theta)^T W^{-1} (y - X\theta)$$

WEIGHTS: $$W = \mathrm{diag}(\sigma_0^2, \ldots, \sigma_T^2)$$

SOLN: $$\theta = (X^T W^{-1} X)^{-1} X W^{-1} y$$



Weighted
Least Squares



Linear Regression

OUTPUTS (Dependent Variables)

$$y_t = \theta^T x_t + v_t$$

$$v_t \sim \mathcal{N}(0, \sigma_t^2)$$

INPUTS (Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_n \end{bmatrix}$$

$$y = X\theta + v$$

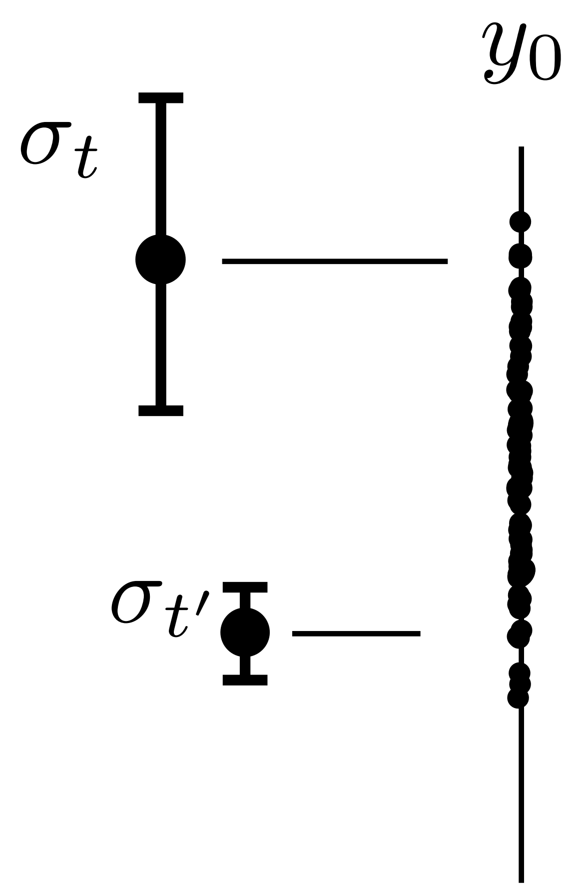
COST: $$\min \quad (y - X\theta)^T W^{-1} (y - X\theta) \quad + (\theta - \mu)^T Q^{-1} (\theta - \mu)$$

Prior on

$$\theta \sim \mathcal{N}(\mu, Q)$$

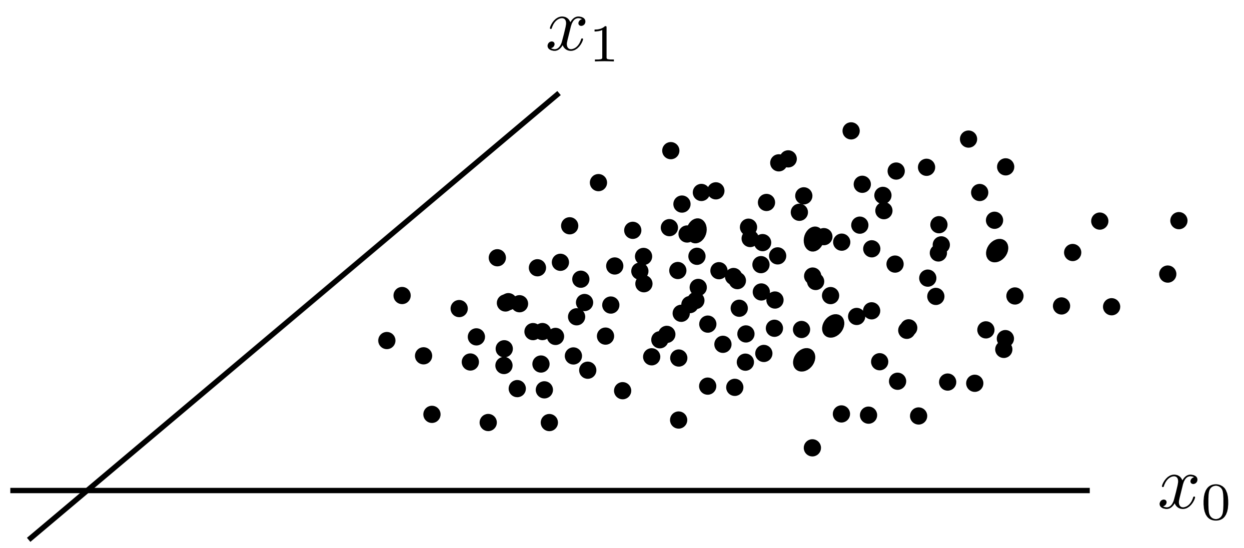
WEIGHTS: $$W = \mathrm{diag}(\sigma_0^2, \ldots, \sigma_T^2)$$

SOLN: $$\theta = \left(X^T W^{-1} X + Q^{-1}\right)^{-1} \left(X^T W^{-1} y + Q^{-1} \mu\right)$$



Bayesian Ridge Regression

$$Q = \lambda^{-1} I$$



Linear Regression

OUTPUTS
(Dependent Variables)

$$y_t = \theta^T x_t + v_t$$

$$v_t \sim \mathcal{N}(0, \sigma_t^2)$$

INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_n \end{bmatrix}$$

$$y = X\theta + v$$

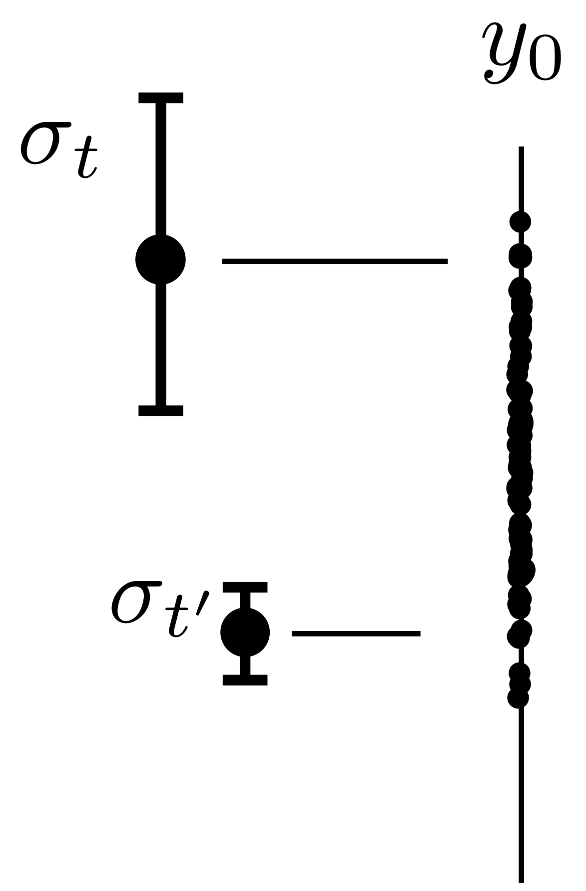
COST: $$\min \quad (y - X\theta)^T W^{-1} (y - X\theta) \quad + (\theta - \mu)^T Q^{-1} (\theta - \mu)$$

Prior on

$$\theta \sim \mathcal{N}(\mu, Q)$$

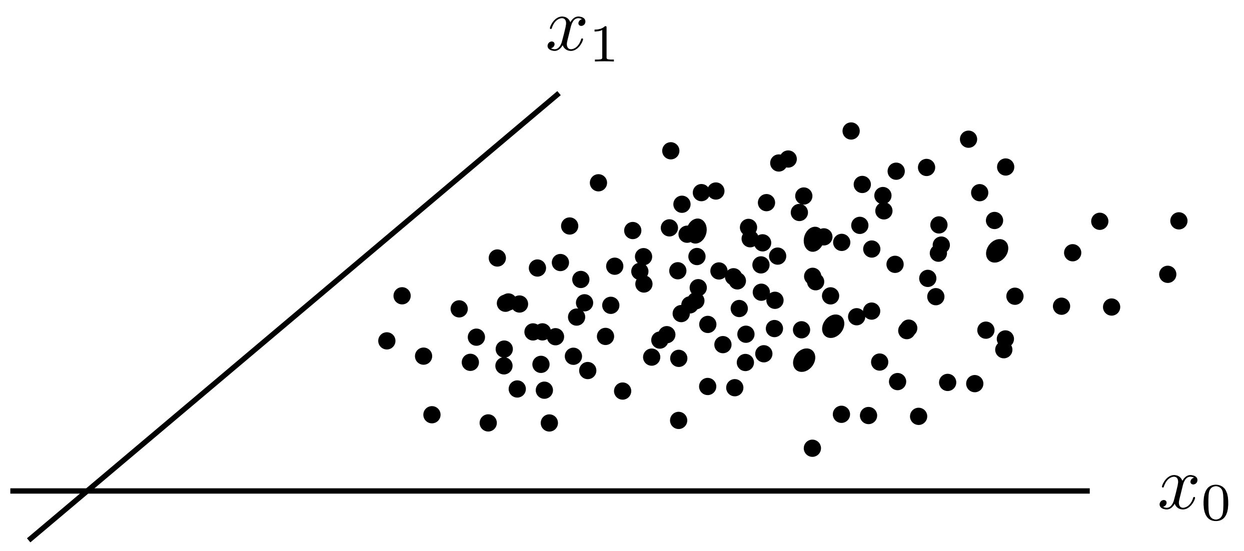
WEIGHTS: $$W = \operatorname{diag}(\sigma_0^2, \ldots, \sigma_T^2)$$

SOLN: $$\theta = \left(X^T W^{-1} X + Q^{-1}\right)^{-1} \left(X^T W^{-1} y + Q^{-1} \mu\right)$$



Automatic Relevance Determination (ARD)

$$Q = \operatorname{diag}(\lambda_0, \ldots, \lambda_n)^{-1}$$



Linear Regression

OUTPUTS
(Dependent Variables)

$$y_t = \theta^T x_t + v_t$$

$$v_t \sim \mathcal{N}(0, \sigma_t^2)$$

INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Longleftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_n \end{bmatrix}$$

$$y = X\theta + v$$

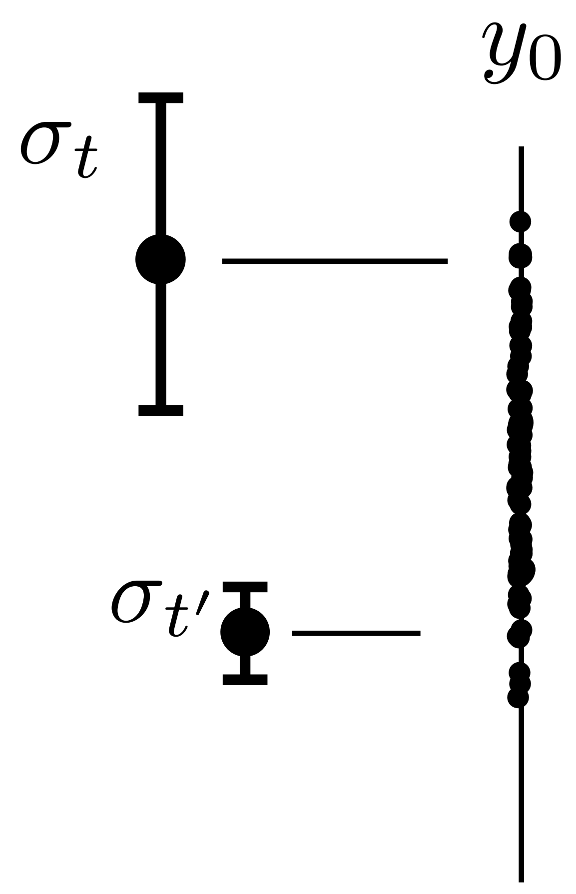
COST: $$\min \quad (y - X\theta)^T W^{-1} (y - X\theta) \quad + (\theta - \mu)^T Q^{-1} (\theta - \mu)$$

Prior on

$$\theta \sim \mathcal{N}(\mu, Q)$$

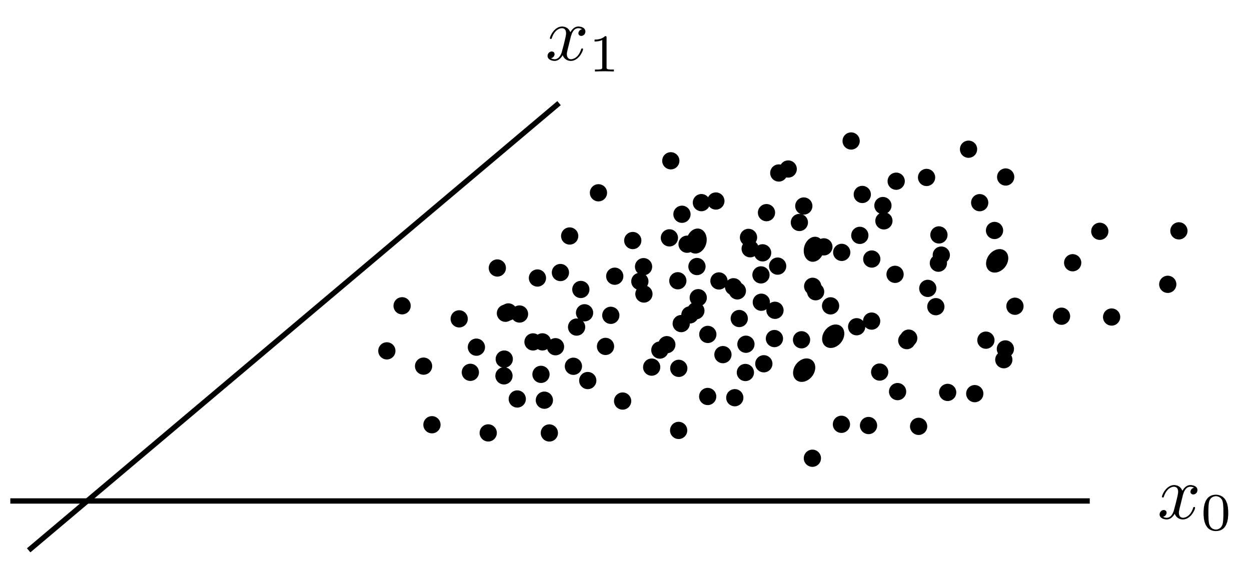
WEIGHTS: $$W = \operatorname{diag}(\sigma_0^2, \ldots, \sigma_T^2)$$

SOLN: $$\theta = \left(X^T W^{-1} X + Q^{-1}\right)^{-1} \left(X^T W^{-1} y + Q^{-1}\mu\right)$$



General Bayesian Regression

$$Q \succ 0$$



Linear Regression

OUTPUTS
(Dependent Variables)

$$y_t = \theta^T x_t$$

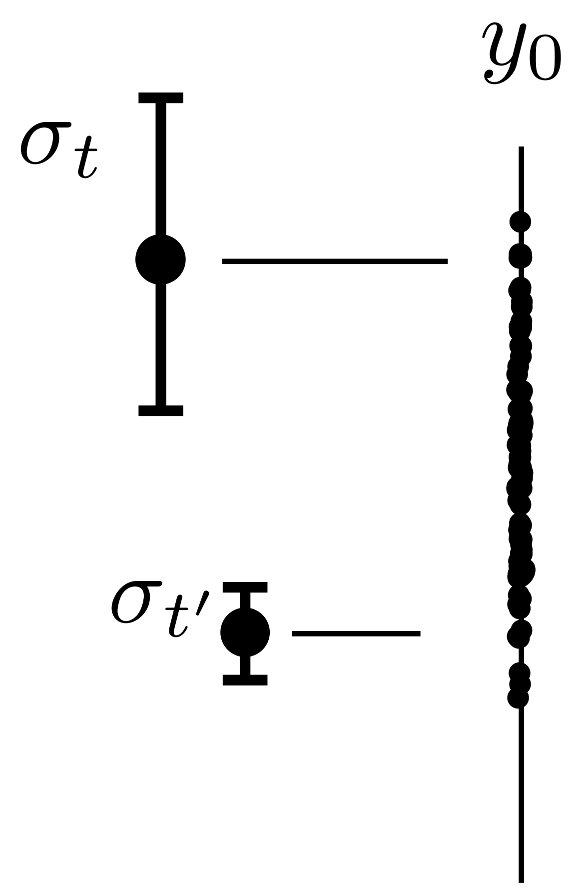
INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_n \end{bmatrix}$$

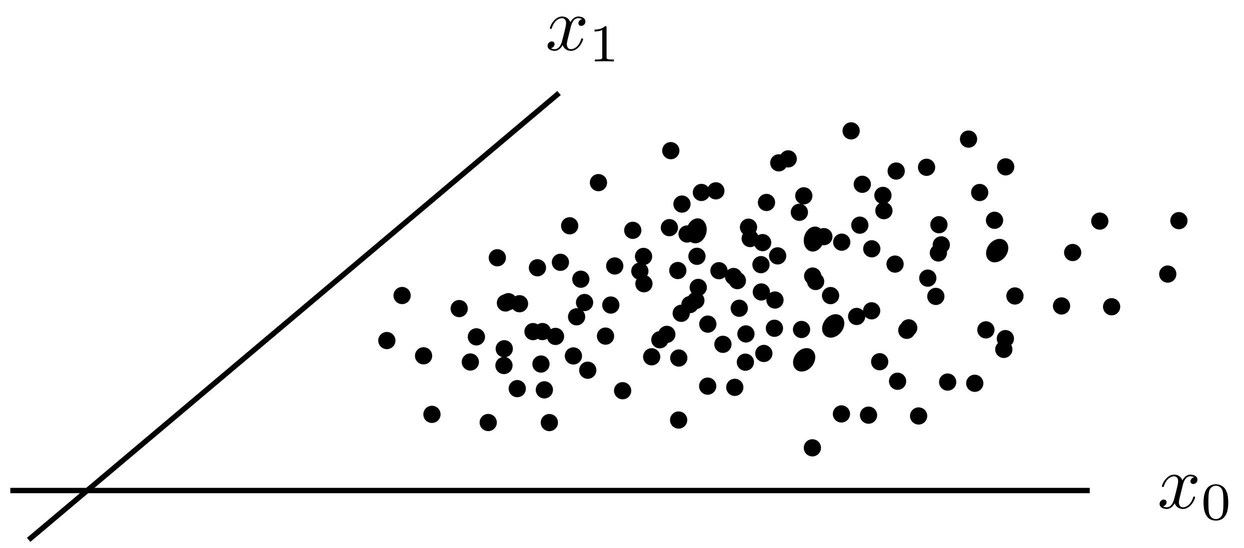
$$y = X\theta$$

COST: $$\min \; \left|\left|y - X\theta\right|\right|^2 \; + \rho(\theta)$$

REGULARIZATION



Ridge Regression $$\rho(\theta) \propto \left|\left|\theta\right|\right|_2^2$$ **"small θ"**



Linear Regression

$$y_t = \theta^T x_t$$

OUTPUTS
(Dependent Variables)

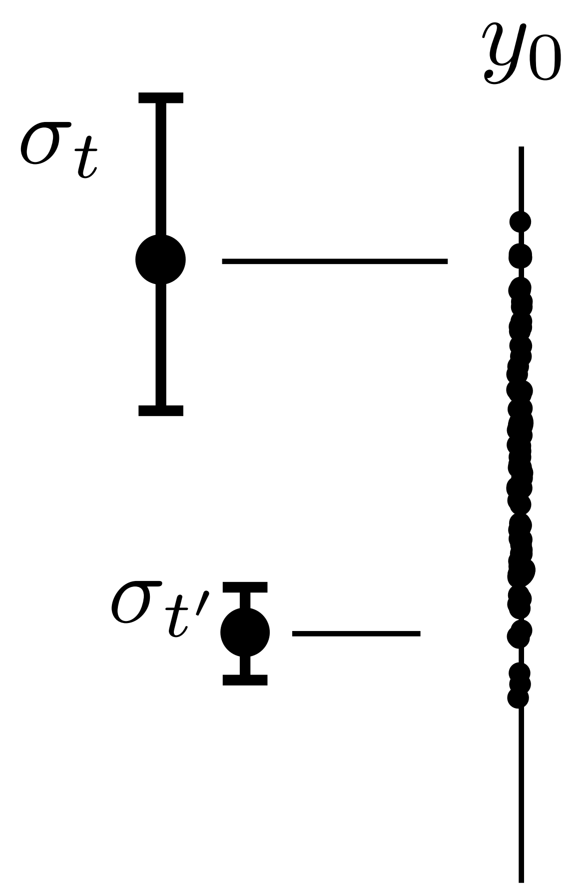
INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix} \begin{bmatrix} 0 \\ \theta_1 \\ \vdots \\ 0 \\ \theta_n \end{bmatrix}$$

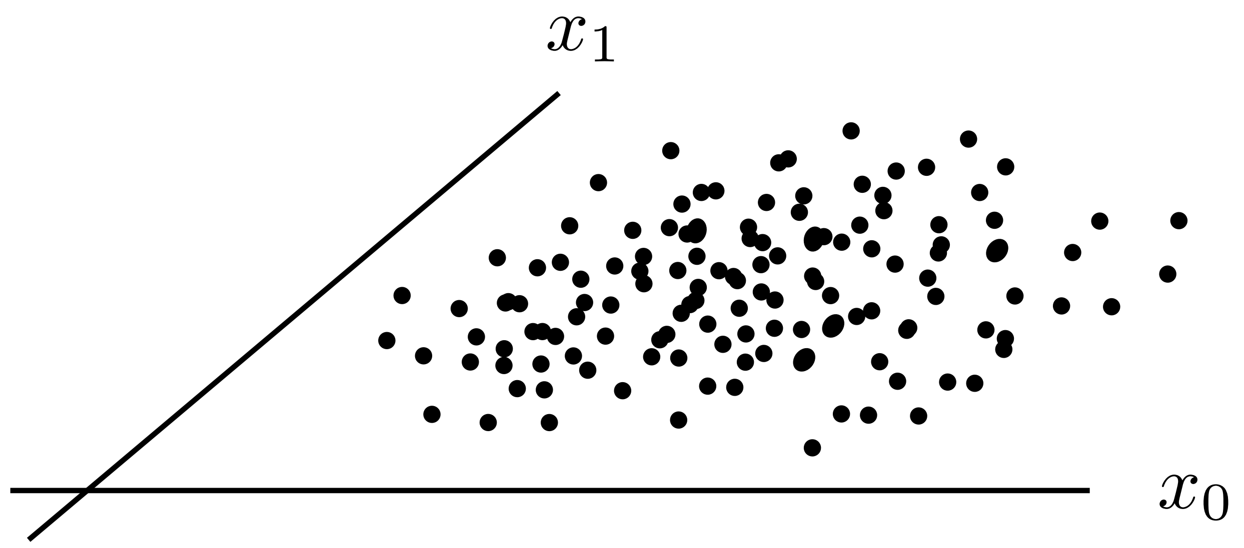
$$y = X\theta$$

COST: $$\min \; \left|\left| y - X\theta \right|\right|^2 + \rho(\theta)$$

REGULARIZATION



Ridge Regression	$\rho(\theta) \propto \left|\left|\theta\right|\right|_2^2$	**“small θ”**
L1-LASSO	$\rho(\theta) \propto \left|\left|\theta\right|\right|_1$	**“sparsity”**



Linear Regression

$$y_t = \theta^T x_t$$

OUTPUTS
(Dependent Variables)

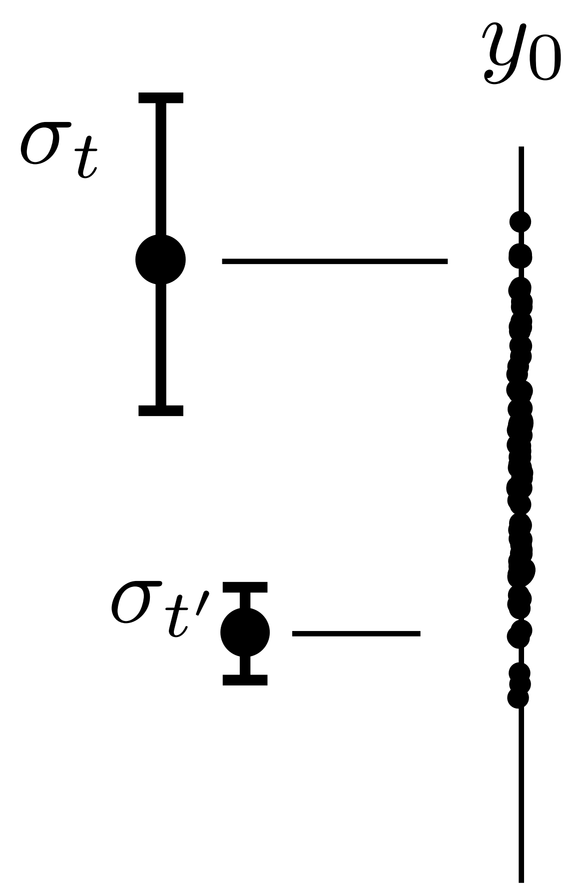
INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix} \begin{bmatrix} 0 \\ \theta_1 \\ \vdots \\ 0 \\ \theta_n \end{bmatrix}$$

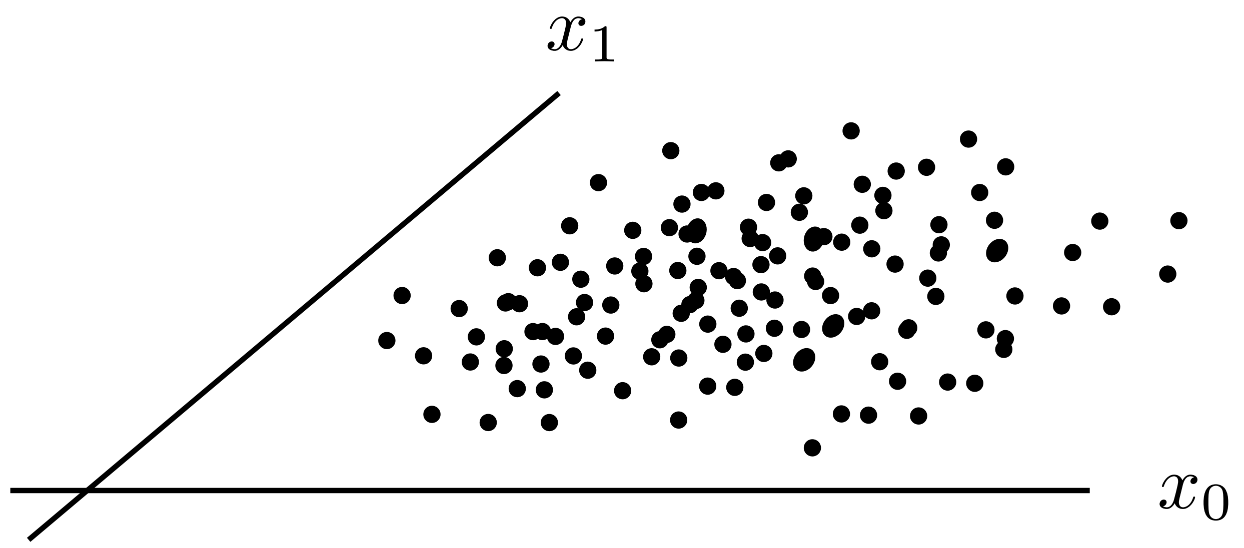
$$y = X\theta$$

COST: $$\min \; \left|\left|y - X\theta\right|\right|^2 + \rho(\theta)$$

REGULARIZATION



Ridge Regression	$\rho(\theta) \propto \left|\left|\theta\right|\right|_2^2$	**"small θ"**
L1-LASSO	$\rho(\theta) \propto \left|\left|\theta\right|\right|_1$	**"sparsity"**
Orthogonal Matching Pursuit (OMP)		**"Sparsity limit"**



Linear Regression

$$y_t = \theta^T x_t$$

OUTPUTS
(Dependent Variables)

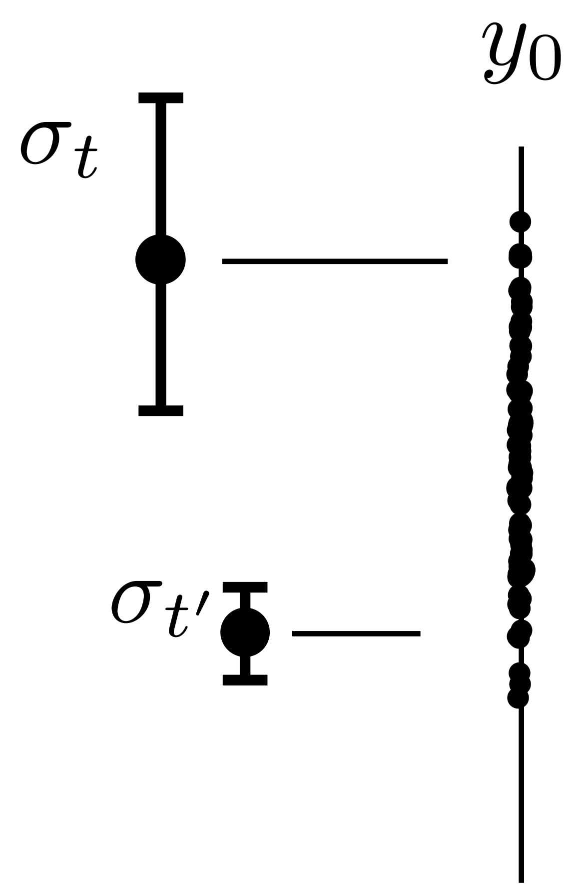
INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix} \begin{bmatrix} 0 \\ \theta_1 \\ \vdots \\ 0 \\ \theta_n \end{bmatrix}$$

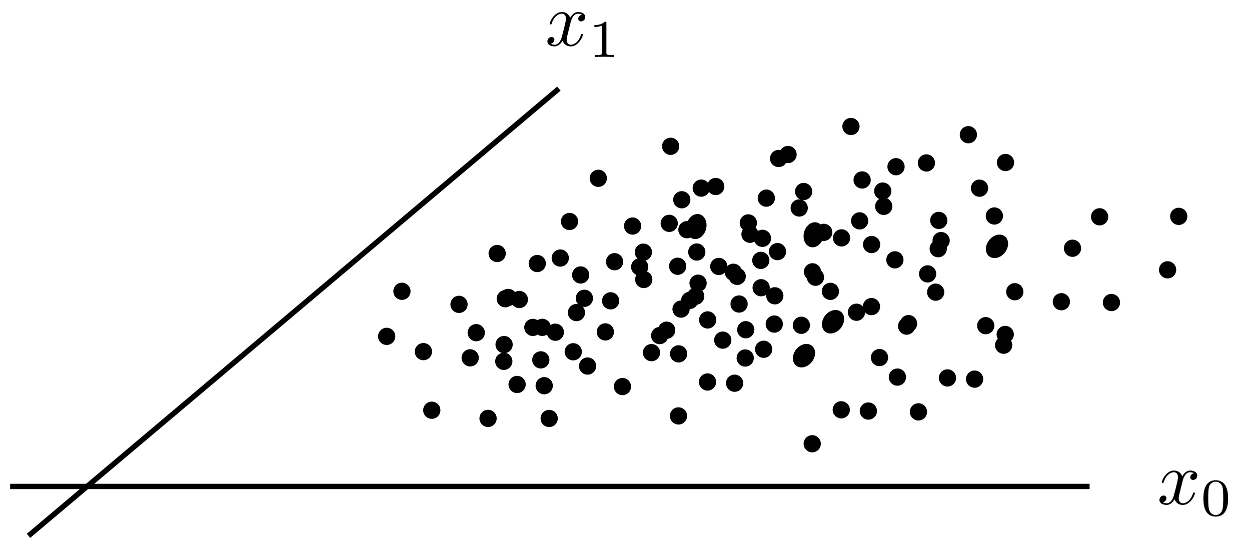
$$y = X\theta$$

COST: $$\min \; \left|\left|y - X\theta\right|\right|^2 + \rho(\theta)$$

REGULARIZATION



Ridge Regression	$\rho(\theta) \propto \left|\left|\theta\right|\right|_2^2$	**"small θ "**
L1-LASSO	$\rho(\theta) \propto \left|\left|\theta\right|\right|_1$	**"sparsity"**
Elastic-Net	$\rho(\theta) \propto \beta\left|\left|\theta\right|\right|_1 + \frac{1}{2}(1-\beta)\left|\left|\theta\right|\right|_2^2$	**"Mix of L1 and L2"**



Linear Regression

OUTPUTS
(Dependent Variables)

$$y_t = \theta^T x_t$$

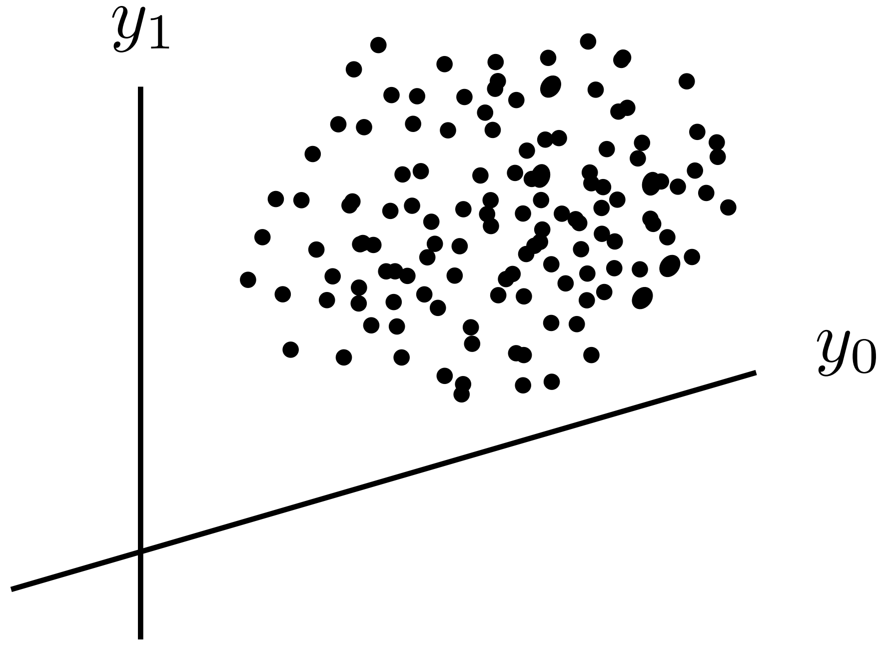
INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix} \begin{bmatrix} 0 & \cdots & 0 \\ \theta_{10} & \cdots & \theta_{1m} \\ \vdots & & \vdots \\ 0 & \cdots & 0 \\ \theta_{n0} & \cdots & \theta_{nm} \end{bmatrix}$$

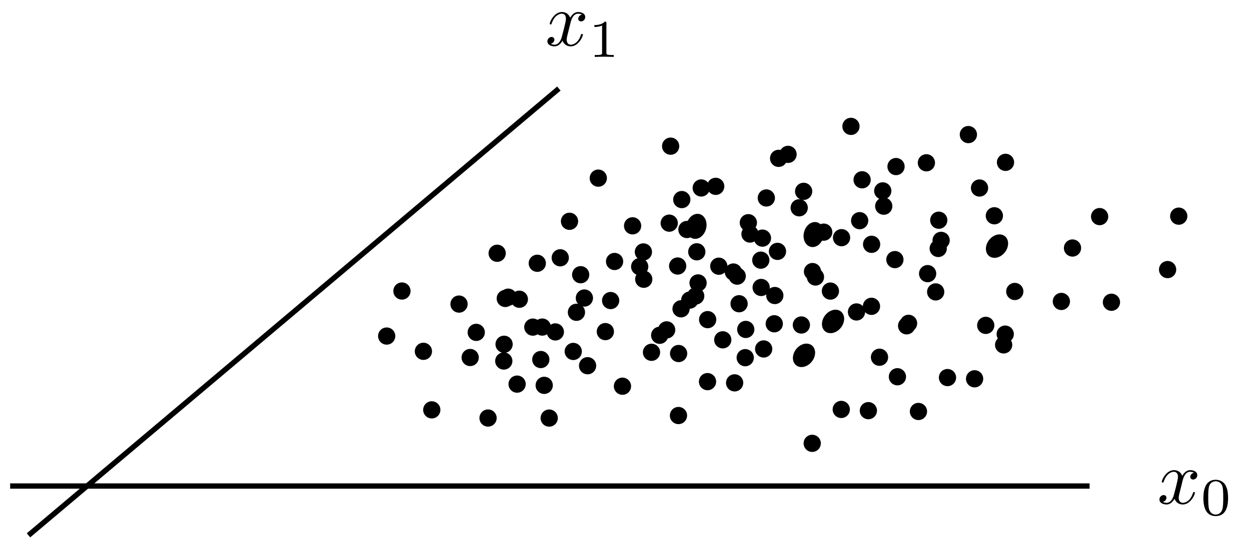
$$y = X\theta$$

COST: $$\min \; \left|\left| y - X\theta \right|\right|_{\text{FRO}}^2 + \rho(\theta)$$

REGULARIZATION



Ridge Regression $\rho(\theta) \propto \left|\left|\theta\right|\right|_{\text{FRO}}^2$ **“small θ ”**



Linear Regression

OUTPUTS
(Dependent Variables)

$$y_t = \theta^T x_t$$

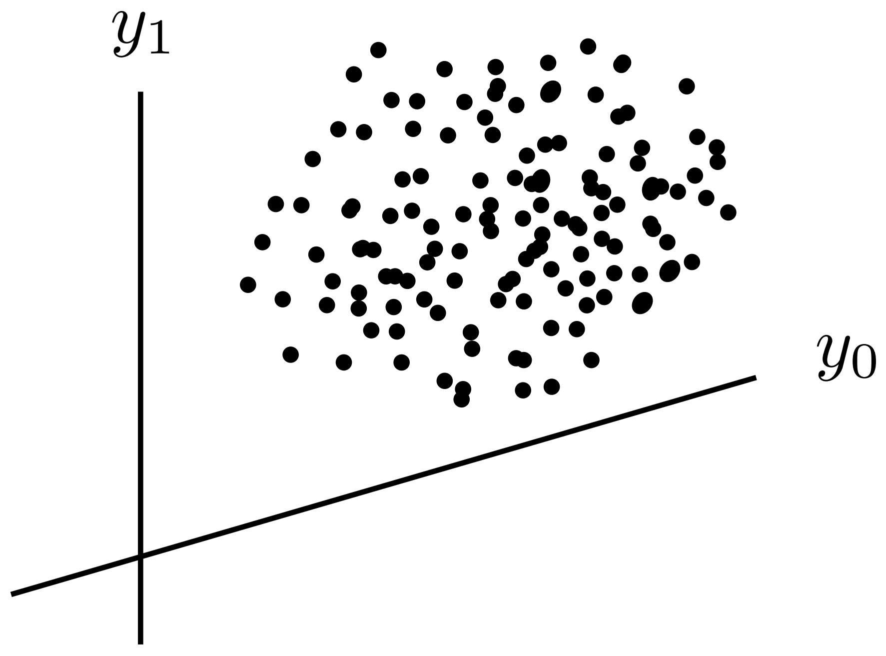
INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix} \begin{bmatrix} 0 & \cdots & 0 \\ \theta_{10} & \cdots & \theta_{1m} \\ \vdots & & \vdots \\ 0 & \cdots & 0 \\ \theta_{n0} & \cdots & \theta_{nm} \end{bmatrix}$$

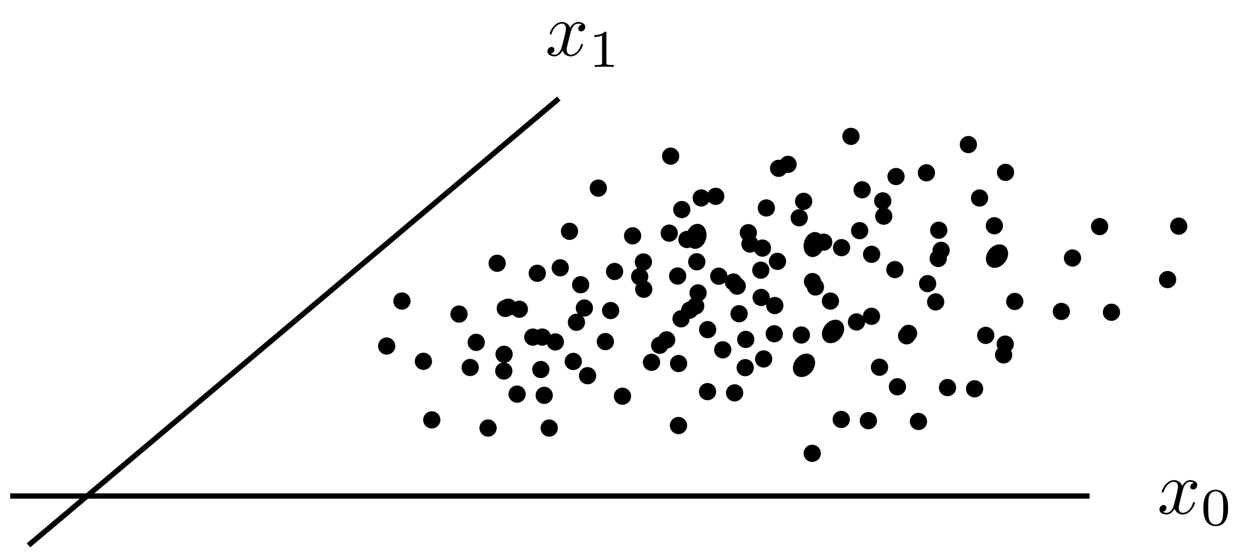
$$y = X\theta$$

COST: $$\min \; \left|\left|y - X\theta\right|\right|^2_{\mathrm{FRO}} + \rho(\theta)$$

REGULARIZATION



Ridge Regression	$\rho(\theta) \propto \left|\left|\theta\right|\right|^2_{\mathrm{FRO}}$	**"small θ "**
Multi-task LASSO	$\rho(\theta) \propto \sum_i \left|\left|\theta_{i:}\right|\right|$	**"Sparsity across tasks"**



Linear Regression

OUTPUTS
(Dependent Variables)

$$y_t = \theta^T x_t$$

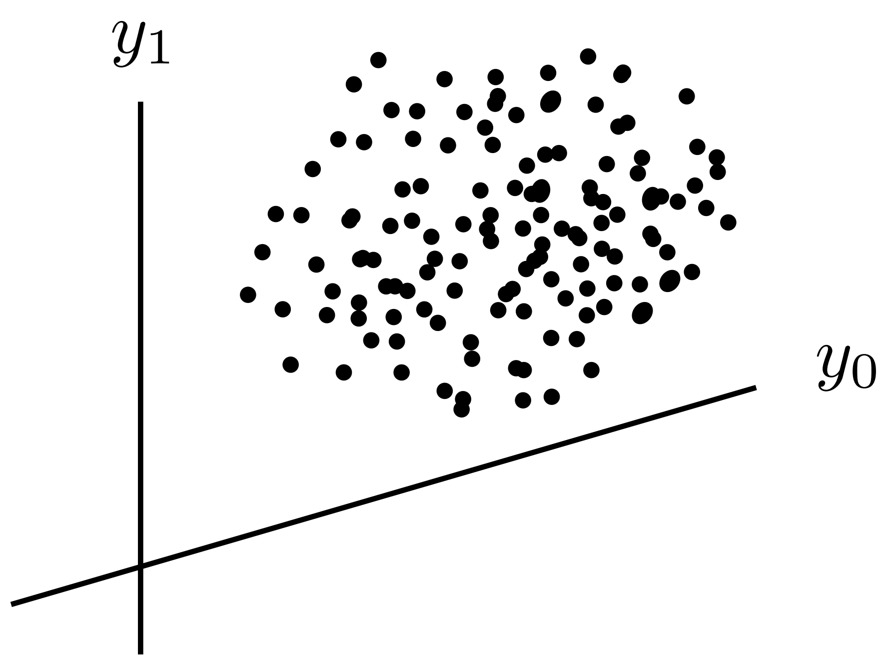
INPUTS
(Independent Variables)

$$\begin{bmatrix} \gamma_{00} & \cdots & \gamma_{0m'} & y_{00} & \cdots & y_{0m} \\ \gamma_{10} & \cdots & \gamma_{1m'} & y_{10} & \cdots & y_{1m} \\ \gamma_{20} & \cdots & \gamma_{2m'} & y_{20} & \cdots & y_{2m} \\ \gamma_{30} & \cdots & \gamma_{3m'} & y_{30} & \cdots & y_{3m} \\ \gamma_{40} & \cdots & \gamma_{4m'} & y_{40} & \cdots & y_{4m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{T0} & \cdots & \gamma_{Tm'} & y_{T0} & \cdots & y_{Tm} \end{bmatrix} \Leftarrow \begin{bmatrix} x_{00} & \cdots & x_{0n} \\ x_{10} & \cdots & x_{1n} \\ x_{20} & \cdots & x_{2n} \\ x_{30} & \cdots & x_{3n} \\ x_{40} & \cdots & x_{4n} \\ \vdots & & \vdots \\ x_{T0} & \cdots & x_{Tn} \end{bmatrix} \begin{bmatrix} 0 & \cdots & 0 \\ \theta_{10} & \cdots & \theta_{1m} \\ \vdots & & \vdots \\ 0 & \cdots & 0 \\ \theta_{n0} & \cdots & \theta_{nm} \end{bmatrix}$$

$$y = X\theta$$

COST: $$\min \; \left|\left|y - X\theta\right|\right|^2_{\text{FRO}} + \rho(\theta)$$

REGULARIZATION



Ridge Regression	$\rho(\theta) \propto \left|\left|\theta\right|\right|^2_{\text{FRO}}$	**"small θ"**
Multi-task LASSO	$\rho(\theta) \propto \sum_i \left|\left|\theta_{i:}\right|\right|$	**"Sparsity across tasks"**
Multi-Task Elastic-Net	$\rho(\theta) \propto \beta \sum_i \left|\left|\theta_{i:}\right|\right|_2 + \frac{1}{2}(1-\beta)\left|\left|\theta\right|\right|^2_{\text{FRO}}$	**"Mixed L1 & L2 across tasks"**