MAT 170 (Spring 2021)

— optimization methods in "data & decisions".

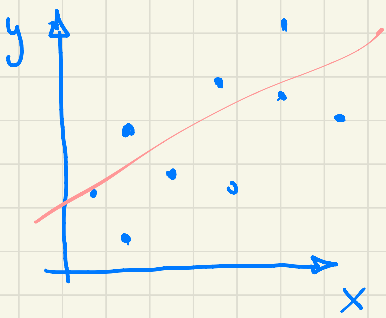
Regression:

Given inputs $x \in \mathbb{R}^m$, $x^1, \dots, x^k$

Given outputs $y \in \mathbb{R}^n$, $y^1, \dots, y^k$

find a simple relationship..

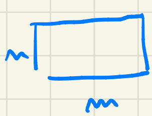
$m = 1$, $n = 1$



Theory: This might be a linear relationship

$$y = f(x) = A x + b$$

$\in \mathbb{R}^n$



Loss function: Measure the "total error" between proposed theory & observed data.

Least squares loss:

$$L(A, b) = \sum \text{square errors}$$

$$= \frac{1}{k} \sum_{i=1}^{k} \| y^i - f(x^i) \|^2$$

$$= \frac{1}{k} \sum_{i=1}^{k} \| y^i - (Ax^i + b) \|^2$$

ℓ^1 regression — sum of norms loss

$$L(A, b) = \frac{1}{k} \sum_{i=1}^{k} \| y^i - f(x^i) \|$$

in particular for <u>scalar</u> outputs $(m = 1,\ y^i \in \mathbb{R})$

$$= \frac{1}{k} \sum_{i=1}^{k} | y^i - f(x^i) |.$$

Formulate it as an opt. problem

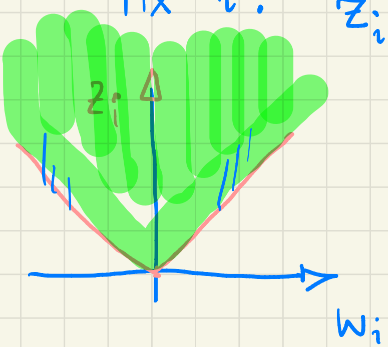
$$\min \quad \frac{1}{k} \sum_{i=1}^{k} | y^i - f(x^i) |$$

nonlinear objective

$$\text{s.t.} \quad A \in \mathbb{R}^{1 \times m}, \quad b \in \mathbb{R}$$

Reformulate this nonlinear objective function so that we obtain an LP.

Fix i.



$$z_i = |y^i - f(x^i)| = |\underbrace{y^i - (Ax^i + b)}_{w_i}|$$

Reformulate:

$$\min \sum_{i=1}^{k} z_i$$ linear objective

s.t. $\boxed{z_i = |w_i|}$ nonlinear constraint

$w_i = y^i - (Ax^i + b)$

Claim: I can relax the problem without changing the optimal solution by replacing $\boxed{=}$ by $\geq$.

Optimality argument:

No optimal solution (A, b, w, z) to this extended formulation can have "$>$".

This convexifies the feasible region.

("epigraphical reformulation")

Rewrite the constraint

$$z_i \geq |w_i|$$

by two linear inequalities

$$z_i \geq w_i$$
$$z_i \geq -w_i.$$

$\Rightarrow$ LP reformulation of ℓ^1 regression:

$$\min \sum z_i$$
$$\text{s.t.} \quad z_i \geq w_i$$
$$z_i \geq -w_i.$$
$$w_i = y^i - (Ax^i + b)$$
$$A \in \mathbb{R}^{1 \times m}, \quad b \in \mathbb{R}$$

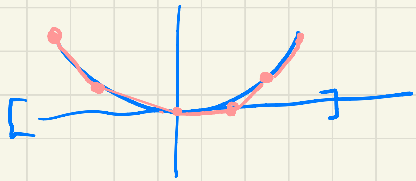
(Variables w_i can be eliminated by plugging the RHS back into inequalities)

$\Rightarrow$ Works for any convex piecewise linear function.

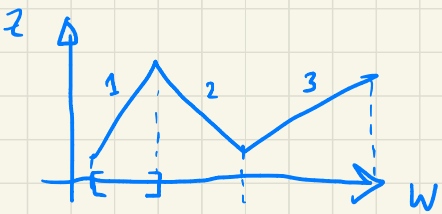
What if we have a convex function that is not piecewise linear?

MICO

Approximate using convex piecewise linear functions



What can we do if we have piecewise linear but nonconvex functions?



Use MILP modeling.

Introduce binary variables to select the piece in which w lies.

$$b_i = \begin{cases} 1 & \text{if } w \text{ lies in } i\text{-th interval} \\ 0 & \text{otherwise} \end{cases}$$

"w should lie in exactly one of the pieces"

$$b_1 + b_2 + b_3 = 1$$