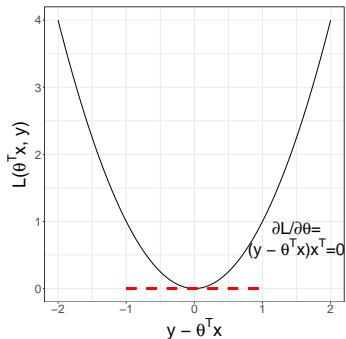


# Introduction to Machine Learning

## Supervised Regression In a Nutshell



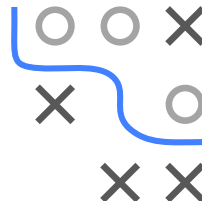
### Learning goals

- Understand basic concept of regressors
- Understand difference between L1 and L2 Loss
- Know basic idea of OLS estimator

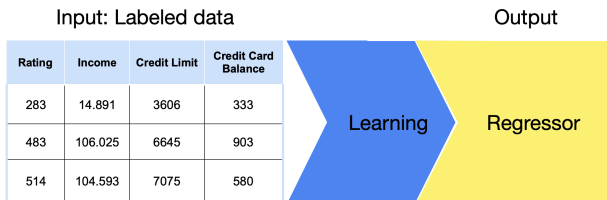
# LINEAR REGRESSION TASKS

- Learn linear combination of features for predicting the target variable
- Find best parameters of the model by training w.r.t. a loss function

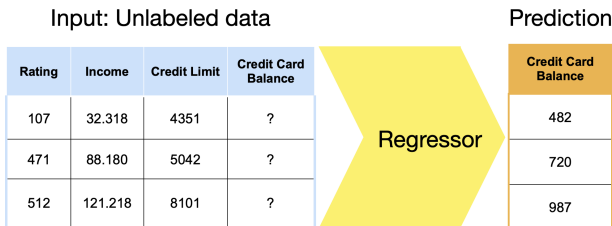
$$\text{CreditBalance} = \theta_0 + \theta_1 \text{Rating} + \theta_2 \text{Income} + \theta_3 \text{CreditLimit}$$



Training

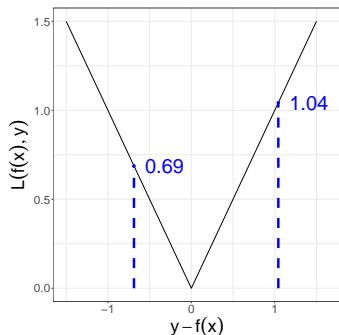
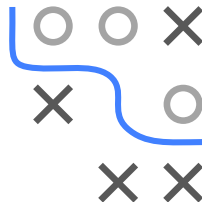


Prediction

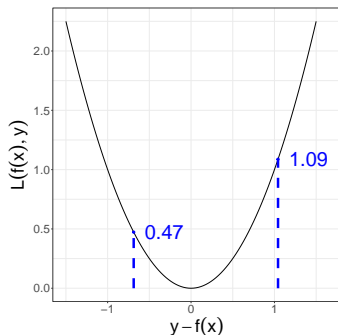


# LINEAR MODELS: L1 VS L2 LOSS

Loss can be characterized as a function of residuals  $r = y - f(\mathbf{x})$



- **L1** penalizes the **absolute** value of residuals
- $L(r) = |r|$
- Robust to outliers

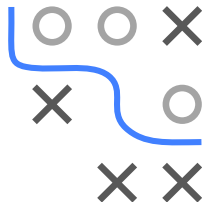


- **L2** penalizes the **quadratic** value of residuals
- $L(r) = r^2$
- Easier to optimize

# LINEAR MODELS: L1 VS L2 LOSS

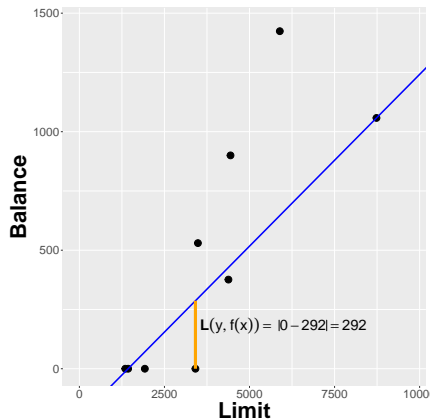
- **L1** Loss is not differentiable in  $r = 0$
- Optimal parameters are computed numerically

- **L2** is a smooth function hence it is differentiable everywhere
- Optimal parameters can be computed analytically or numerically

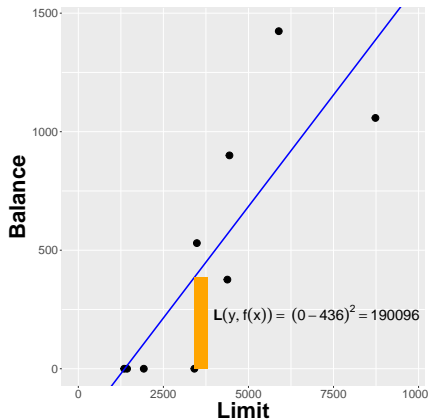


# LINEAR MODELS: L1 VS L2 LOSS

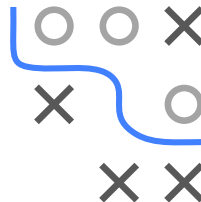
- The parameter values of the best model depend on the loss type



- $\hat{\theta}_{L_1} = 0.14 \rightarrow$  if the Credit Limit increases by 1\$ the Credit Balance increases by 14 Cents



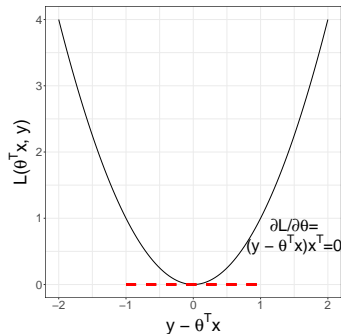
- $\hat{\theta}_{L_2} = 0.19 \rightarrow$  if the Credit Limit increases by 1\$ the Credit Balance increases by 19 Cents



# OLS ESTIMATOR

Ordinary-Least-Squares (OLS) estimator:

- Analytical solution for linear models with L2 loss
- Best parameters can be computed via derivation of the empirical risk
- Solution:  $\hat{\theta} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$



# OLS ESTIMATOR

Components of **OLS** estimator:

- **X**: Features + extra column for intercept
- **y**: Label vector

 $X$ 

| Intercept | Rating | Income  | Credit Limit |
|-----------|--------|---------|--------------|
| 1         | 283    | 14.891  | 3606         |
| 1         | 483    | 106.025 | 6645         |
| 1         | 514    | 104.593 | 7075         |

 $y$ 

| Credit Card Balance |
|---------------------|
| 333                 |
| 903                 |
| 580                 |

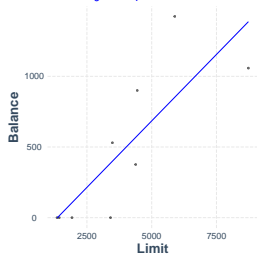


# POLYNOMIAL REGRESSION

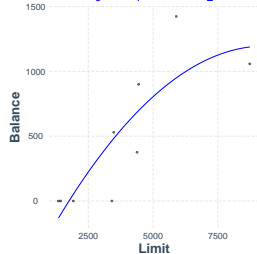
- Adding polynomial terms to the linear combination leads to more flexible regression functions
- Too high degrees can lead to overfitting



$$\text{Balance} = \theta_0 + \theta_1 \text{Limit}$$



$$\text{Balance} = \theta_0 + \theta_1 \text{Limit} + \theta_2 \text{Limit}^2$$



$$\text{Balance} = \theta_0 + \theta_1 \text{Limit} + \theta_2 \text{Limit}^2 + \theta_3 \text{Limit}^3$$

