# Linear Regression with Multiple Variables:

Linear Regression with Multiple Variables, often referred to as **Multiple Linear Regression** or **Multivariate Linear Regression**, is an extension of simple linear regression that involves predicting a target variable based on two or more independent variables (features). It's a fundamental statistical technique used in various fields, including statistics, economics, engineering, and machine learning.

Mean normalization

$$x_1 = \frac{x_1 - \mu_1}{\underbrace{2000-300}_{\text{max}-\text{min}}} \qquad x_2 = \frac{x_2 - \mu_2}{\underbrace{5-0}_{\text{max}-\text{min}}}$$

$$-0.18 \le x_1 \le 0.82 \qquad -0.46 \le x_2 \le 0.54$$

## Key Concepts:

1. **Equation:** The multiple linear regression equation takes the form:

$$\mathbf{y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon}$$

**where:**

- **y:** The target (dependent) variable you want to predict.
- **$x_1, x_2, \ldots, x_n$:** The independent variables (features) that influence the target.
- **$\beta_0, \beta_1, \beta_2, \ldots, \beta_n$:** Coefficients (weights) that represent the impact of each feature.
- **ε:** The error term, accounting for unexplained variability.

2. **Matrix Notation:** Multiple linear regression can also be represented using matrix notation:

$$\mathbf{Y = X\beta + \varepsilon}$$

**where:**

- **Y:** A column vector of target values.
- **X:** A matrix of feature values.
- **β:** A column vector of coefficients.
- **ε:** A column vector of error terms.

Gradient descent

One feature

repeat {

$$w = w - \alpha \underbrace{\frac{1}{m}\sum_{i=1}^{m}\left(f_{w,b}(x^{(i)}) - y^{(i)}\right)x^{(i)}}_{\frac{\partial}{\partial w}J(w,b)}$$

$$b = b - \alpha \frac{1}{m}\sum_{i=1}^{m}\left(f_{w,b}(x^{(i)}) - y^{(i)}\right)$$

simultaneously update $w$, $b$

}

$n$ features ($n \ge 2$)

repeat {

$j=1$

$$w_1 = w_1 - \alpha \underbrace{\frac{1}{m}\sum_{i=1}^{m}\left(f_{\vec{w},b}(\vec{x}^{(i)}) - y^{(i)}\right)x_1^{(i)}}_{\frac{\partial}{\partial w_1}J(\vec{w},b)}$$

⋮

$j=n$

$$w_n = w_n - \alpha \frac{1}{m}\sum_{i=1}^{m}\left(f_{\vec{w},b}(\vec{x}^{(i)}) - y^{(i)}\right)x_n^{(i)}$$

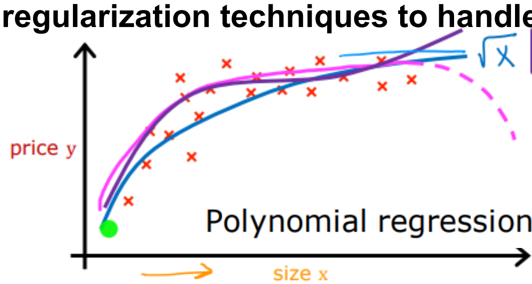
$$b = b - \alpha \frac{1}{m}\sum_{i=1}^{m}\left(f_{\vec{w},b}(\vec{x}^{(i)}) - y^{(i)}\right)$$

simultaneously update $w_j$ (for $j = 1, \cdots, n$) and $b$

}

3. **Objective:** The goal is to find the coefficients (β values) that minimize the sum of squared differences between the predicted values and the actual target values.
4. **Assumptions:** Multiple linear regression assumes several key assumptions, including linearity, independence of errors, constant variance (homoscedasticity), and normality of errors.
5. **Coefficient Interpretation:** The coefficients (β values) indicate how much the target variable changes for a one-unit change in the corresponding feature while keeping other features constant.
6. **Feature Selection:** It's important to select relevant features and avoid multicollinearity (high correlation between features), as it can affect coefficient interpretation and model performance.
7. **Model Evaluation:** Common evaluation metrics include the coefficient of determination ($R^2$), adjusted $R^2$, and root mean squared error (RMSE), among others.

**Multiple Linear Regression can be extended to more complex models, such as Polynomial Regression, Ridge Regression, Lasso Regression, and Elastic Net Regression, which introduce regularization techniques to handle overfitting and improve model generalization.**




$$\boxed{f_{\vec{w},b}(x) = w_1 x + w_2\sqrt{x} + b}$$

size √size

what features to use?
↳ course 2

ZAGAZIG Student Branch