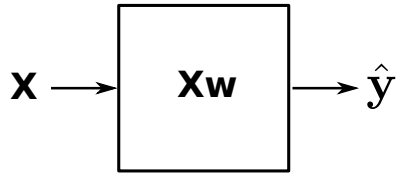


- 29 Example (Linear Regression With More Than One Feature)
Linear regression networks with d features



can be expressed in matrix form as the function

$$\mathbf{f}(\mathbf{X}) = \mathbf{X}\mathbf{w} = \hat{\mathbf{y}}$$

where

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1d} & 1 \\ x_{21} & x_{22} & \dots & x_{2d} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nd} & 1 \end{pmatrix}, \mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ b \end{pmatrix}, \hat{\mathbf{y}} = \begin{pmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{pmatrix}$$

Columns j , $j = 1, 2, \dots, d$, of the data matrix \mathbf{X} contain the n values of the j th feature. w_j is the network weight of the j th feature. The last column corresponds to the bias and can be interpreted as a special feature whose value is always equal to 1.

- 30 Example (Linear Regression for Red Wine)
- Load the data set `wine_quality_red.csv` into a dataframe. Assume wine quality is the feature to be predicted. How many input features does the data set contain? How many data points?
 - Determine the feature weights and bias that minimize MSE. Hint: Add a column of 1's to the data matrix \mathbf{X} to represent the bias. Then solve the normal equations to obtain the optimal weights and bias.
 - Compare RMSE of the linear regression network with the RMSE of a simple bias network.

- 31 Example (Red Wine Training vs Testing RMSE)
Shuffle then split the red wine dataset into two datasets, 80% for training and 20% for testing. Train on the training set and compute RMSE for training and testing sets. Which do you expect to be larger training RMSE or testing RMSE?

- 32 Definition (One-Hot-Encoding)

One-Hot-Encoding is a procedure that converts categorical data into numerical data that can be used as input to machine learning algorithms. The converted data can be interpreted as a probability distribution.

- 33 Example (One-Hot-Encoding Abalone)
Data Source⁴

Abalone is a type of snail. How accurately can the age of abalone be predicted from physical measurements? Data on the features listed in the table below is given in the file `Abalone.csv`.



feature	measurement	description
sex	M, F, I	male, female or infant
length	mm	longest shell measurement
diameter	mm	perpendicular to length
height	mm	with meat in shell
weight_whole	gms	whole abalone
weight_shucked	gms	weight of meat
weight viscera	gms	gut weight (after bleeding)
weight_shell	gms	after being dried
rings		+1.5 gives the age in years

- 34 Definition (Feature Engineering)

Feature engineering is a process of creating new features from old ones in an effort to improve the performance of a machine learning al-

⁴Marine Resources Division, Marine Research Laboratories-Taroona, Department of Primary Industry and Fisheries, GPO Box 619F, Hobart, Tasmania 7001, Australia