Linear Functions

Function notation: $f: R^n \rightarrow R$ is a function that maps real n-vectors to real numbers.

- If x is a vector, then $f(x)$ is a scalar. (x is the argument)

$$f(x) = f(x_1, x_2, x_3 \ldots\ldots x_n)$$

$\Rightarrow$ f satisfies the superposition property if $\boxed{f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)}$

- Such a function that satisfies superposition is called linear.

Inner Product Function: a is an n-vector.

$$\boxed{f(x) = a^T x} = a_1x_1 + a_2x_2 + \ldots\ldots + a_nx_n$$

Proving superposition for the above:

$$f(\alpha x + \beta y) = a^T(\alpha x + \beta y)$$
$$= \alpha a^T x + \beta a^T y$$
$$= \alpha f(x) + \beta f(y)$$

$\therefore$ Inner product is linear function.

* If $f()$ is linear, superposition extends to linear combinations of any number of vectors and not just 2 vectors.

- A function $f: R^n \rightarrow R$ is linear if it satisfies
  - Homogenity: $f(\alpha x) = \alpha f(x)$
  - Additivity: $f(x+y) = f(x) + f(y)$

Inner product representation of linear function:

$\rightarrow$ If a function is linear, it can be expressed as the inner product with some fixed vector.

Ex: We know that $x = x_1e_1 + x_2e_2 \ldots + x_ne_n$ is a linear function.

By multi-term superposition, $\therefore f(x_1e_1 + x_2e_2 + \ldots + x_ne_n)$

$$\Rightarrow x_1f(e_1) + x_2f(e_2) + \ldots + x_nf(e_n)$$
$$= a^T x \quad (\text{where } a = [f(e_1), f(e_2), f(e_3) \ldots, f(e_n)])$$

$f(x) = a^T x$ is unique. There is only one vector 'a' for which $f(x) = a^T x$ holds $\forall x$.

Avg of vector: of an n-vector is $f(x) = (x_1 + x_2 + \ldots + x_n)/n$. $= avg(x) / \bar{x}$

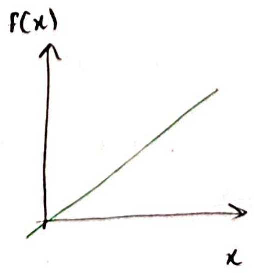
This is a linear function as it can be expressed as $avg(x) = a^T x$.

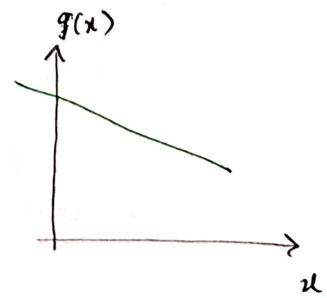
Affine functions: A function that is linear plus a constant $\boxed{f(x) = a^T x + b}$ offset.

a function $f: R^n \rightarrow R$ is affine iff:
$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$
holds $\forall \alpha, \beta$ and $\alpha + \beta = 1$

Affine functions are not always linear.

$y = 4x$ linear $\quad y = 4x + 6$ affine





* Linear Function must pass through origin *

* For linear functions, superposition holds for any coefficients α & β. But for affine fn's it holds when the coefficients sum to one. (i.e, $\alpha + \beta = 1$)

Taylor Approximation:

- suppose you have a function $f : R^n \longrightarrow R$.

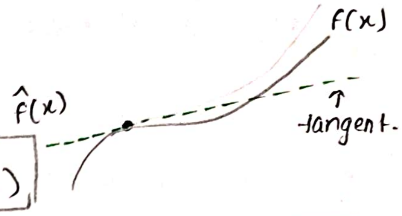
- Differential calculus gives us an organized way to find an approximate affine model.

First order taylor approximation:

$$\boxed{\hat{f}(x) = f(z) + \frac{\partial f}{\partial x_1}(z)(x_1 - z_1) + \ldots\ldots + \frac{\partial f}{\partial x_n}(z)(x_n - z_n)}$$

The ^ means an approximation.

- $\frac{\partial f}{\partial x_i}(z)$ = partial denivitive of f with respect to its ith argument, evaluated at n-vec z.
- $\hat{f}(x)$ is very close to $f(x)$ when x_i are all near z_i
- $\hat{f}$ is an affine function of x.



Inner product representation is: $\boxed{\hat{f}(x) = f(z) + \nabla f(z)^T (x - z)}$

where $\nabla f(x)$ is the gradiant of f at z: $\boxed{\nabla f(z) = \left(\frac{\partial f}{\partial x_1} z, \frac{\partial f}{\partial x_2} z, \ldots \frac{\partial f}{\partial x_n}(z)\right)}$

Ex: $f : R^2 \longrightarrow R$; $f(x) = x_1 + \exp(x_2 - x_1)$; $z = (1, 2)$; find 1st taylor approx.

Ans: $f(z) = 1 + \exp(2-1) = 3.7183$

$$\nabla f(z) = \left(\frac{\partial f(z)}{\partial x_1}, \frac{\partial f(z)}{\partial x_2}\right) = \left(1 + \exp(z_2 - z_1)(-1), \exp(z_2 - z_1)\right)$$

$$= (-1.7183, 2.7183)$$

$$\therefore \hat{f}(x) = f(z) + \nabla f(z)^T (x - z)$$

$$= 3.7183 + (-1.7183, 2.7183)^T (x - (1,2))$$

$$= 3.7183 + -1.7183(x_1 - 1) + 2.7183(x_2 - 2)$$

Regression Model:

The regression model is an affine function of the vector (x).

n-vector $\boxed{\hat{y} = x^T \beta + v}$ scalar regressors

where, y is called dependant variable, outcome or label.

n-vector is called weight vector | coefficient vector.

v is called offset or intercept.

Parameters = β & v