

Week 2.2 Computing Parameters Analytically

Normal Equation

- Normal equation: Method to solve for θ analytically
- $\theta = (X^T X)^{-1} X^T y$

[详解正规方程 \(Normal Equation\)](#)

```
pinv (X'*X)*X'*y
```

- X: design matrix

Examples: $m = 4$.

	Size (feet ²)	Number of bedrooms	Number of floors	Age of home (years)	Price (\$1000)
x_0	x_1	x_2	x_3	x_4	y
1	2104	5	1	45	460
1	1416	3	2	40	232
1	1534	3	2	30	315
1	852	2	1	36	178

$$X = \begin{bmatrix} 1 & 2104 & 5 & 1 & 45 \\ 1 & 1416 & 3 & 2 & 40 \\ 1 & 1534 & 3 & 2 & 30 \\ 1 & 852 & 2 & 1 & 36 \end{bmatrix}$$

$m \times (n+1)$

$$y = \begin{bmatrix} 460 \\ 232 \\ 315 \\ 178 \end{bmatrix}$$

m -dimensional vector

$$\theta = (X^T X)^{-1} X^T y$$

- Gradient Descent vs Normal Equation:

Gradient Descent	Normal Equation
Need to choose alpha	No need to choose alpha
Needs many iterations	No need to iterate
$O(kn^2)$	$O(n^3)$, need to calculate inverse of $X^T X$
Works well when n is large	Slow if n is very large

Normal Equation Noninvertibility

- What if $X^T X$ is non-invertible? (rarely)
- pinv: compute inverse even when invertible
- Causes:

- Redundant features (linearly dependent)
- Too many features (e.g. $m \leq n$)
 - delete or use regularization