**Linear Regresyon Maliyet (Cost) Fonksiyonu**

Oluşturulan Linear Eğrinin doğruluğu hakkında bilgi edinebilmek için maliyet fonksiyonu kullanırız. Aşağıdaki formül ile ifade ederiz;

Bu denklemi karesel hata fonksiyonu olarak adlandırabiliriz. Tahmin edilen ile gerçek değer arasındaki farkı ölçer. Amacımız buradaki değerin küçük elde edilmesidir.

**Gradient Descent ile Parametre Öğrenmesi**

Cost fonksiyonu ile veri değerlerinin modele ne kadar uyup uymadığını ölçebiliriz. Gradient Descent algoritmasını kullanarak parametrelerin yaklaşık değerlerini elde edebiliriz.

Her iterasyonda değerlerini eş zamanlı olarak güncellemeliyiz.

değeri learning rate (öğrenme oranı) olarak bilinir.Bu değerin seçimi oldukça önemlidir. değeri düşük seçilir ise algoritma yavaş çalışır ve optimum değeri geç zamanda bulur. değeri çok yüksek seçilir ise optimum değeri aşabilir ve optimum değere yakınlaşmada başarısız olur.

Gradient Descent algoritmasında sabit değeri sorun teşkil etmez. Algoritma minimum değere yakınlaştığı zaman değeri otomatik olarak küçülür. Bu sayede değerini zaman geçtikçe düşürmemize gerek kalmaz.

**Gradient Descent ile Regresyon**

değerlerinin güncellemesi şu şekilde yapılır.

değerlerinin güncellemesi local minimum değerine yakınlaşmasına kadar devam eder. m değerini eğitim veri setimizdeki kayıt sayısı olarak ifade ederiz.

**Normal Equation ile Parametre Öğrenimi**

Gradient Descent algoritmasına göre daha yavaş çalışır. Herhangi bir şekilde iterasyona gerek duymaz ve değeri kullanılmaz. Genel formül;

X matrisimiz bağımsız değişkenlerden başka bir deyiş ile tahmin edici değişkenlerden oluşur. Bu matrisin ilk sütununa m değeri kadar bir ekleriz. Y matrisimiz ise çıkış değerlerimizdir. R programla dilindeki lm() fonksiyonu bu yöntemi kullanır.

Eğer ifadesinin tersi alınamıyor ise iki sebep yüzünden olabilir bunlar;

* İki tahmin edici sütun arasındaki benzerlik yüksek ise
* Çok sayıda tahmin edici değişken bulunuyor ise

Çözüm olarak tahmin edici değişken sayısını azaltabiliriz veya birbirine doğrusal olarak yüksek seviyede bağımlı olan iki tahmin edici sütundan birini ortadan kaldırabiliriz.