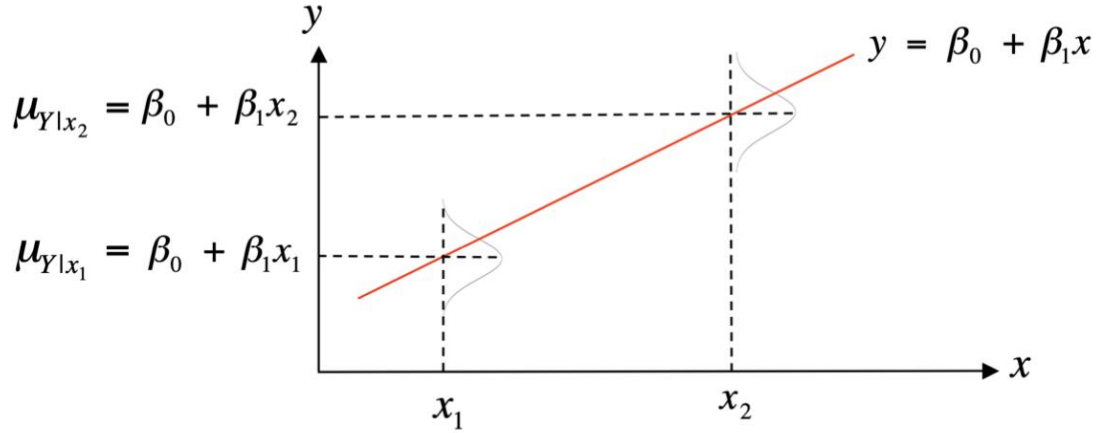


## Grafiksel Yorumlama



Örneğin,  $x$  = boy ve  $y$  = ağırlık ise,  $\mu_{Y|x} = 60$ , popülasyonda 60 inç boyunda olan tüm bireylerin ortalama ağırlığıdır.

## Örnek

Bağımsız değişken yükseklik ( $x$ ) ile bağımlı değişken ağırlık ( $y$ ) arasındaki ilişkinin, gerçek regresyon çizgisine sahip basit bir doğrusal regresyon modeli ile tanımlandığını varsayalım

$$y = 7.5 + 0.5x \text{ and } \sigma = 3$$

- S1:  $\beta_1 = 0.5$ 'in yorumu nedir?

Ağırlık olarak 1 birim artışla ilişkili beklenen yükseklik değişikliği

- S2:  $x = 20$  ise,  $Y$ 'nin beklenen değeri nedir?

$$\mu_{Y|x=20} = 7.5 + 0.5(20) = 17.5$$

- S3:  $x = 20$  ise  $P(Y > 22)$  nedir?

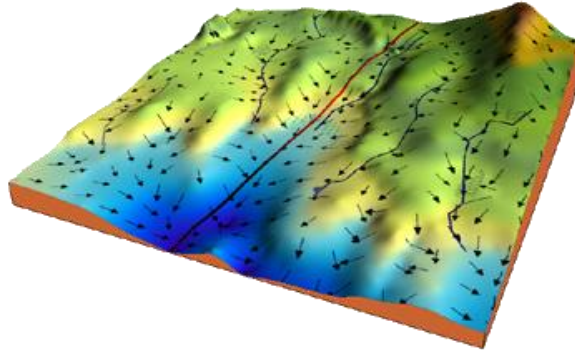
$$P(Y > 22 | x = 20) = P\left(\frac{22 - 17.5}{3}\right) = 1 - \phi(1.5) = 0.067$$

## Dereceli Alçalma (Gradient Descent)

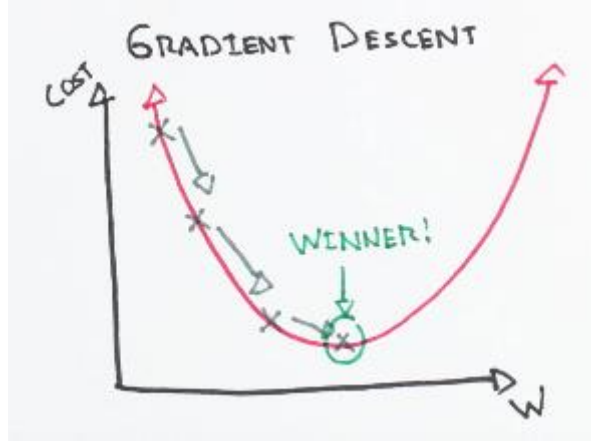
Gradyan inişi, tekrarlı olarak “en dik iniş yönünde” hareket ederek bazı fonksiyonları en aza indirmek için kullanılan bir optimizasyon algoritmasıdır. Makine öğreniminde, modelimizin parametrelerini güncellemek için degrade iniş kullanırız. Parametreler, Doğrusal Regresyondaki katsayıları ve sinir ağlarındaki ağırlıkları ifade eder.

### Giriş

Bir maliyet fonksiyonu bağlamında aşağıdaki 3 boyutlu grafiği düşünün. Hedefimiz sağ üst köşedeki dağdan (yüksek maliyet) sol altta koyu mavi denize (düşük maliyet) geçmektir. Oklar herhangi bir noktadan en dik iniş yönünü (negatif eğim) temsil eder-maliyet fonksiyonunu mümkün olduğunca çabuk azaltan yön.



Dağın tepesinden başlayarak, negatif adımın belirlediği yönde ilk adımımızı yokuş aşağı atıyoruz. Daha sonra negatif degradeyi yeniden hesaplıyoruz (yeni noktamızın koordinatlarından geçerek) ve belirttiği yönde başka bir adım atıyoruz. Grafiğimizin alt kısmına veya artık yokuş aşağı hareket edemediğimiz bir nokta olana kadar yerel minimum olana kadar bu işleme yinelemeli olarak devam ediyoruz.



## Öğrenme Oranı (Learning Rate)

Bu adımların boyutuna öğrenme oranı denir. Yüksek bir öğrenme oranı ile her adımda daha fazla yer kaplayabiliriz, ancak tepenin eğimi sürekli değiştiği için en düşük noktayı aşma riski taşırız. Çok düşük bir öğrenme oranıyla, negatif gradyan yönünde güvenle hareket edebiliriz, çünkü bunu çok sık yeniden hesaplıyoruz. Düşük bir öğrenme oranı daha kesindir, ancak gradyanı hesaplamak zaman alıcıdır, bu nedenle en alt noktaya gelmek çok uzun zaman alacaktır.

## Maliyet Fonksiyonu (Cost Function)

Kayıp Fonksiyonları (Loss Functions), modelimizin belirli bir parametre seti için tahminlerde bulunmada “ne kadar iyi” olduğunu söyler. Maliyet fonksiyonunun kendi eğrisi ve kendi degradeleri vardır. Bu eğrinin eğimi, modeli daha doğru hale getirmek için parametrelerimizi nasıl güncelleyeceğimizi anlatıyor.