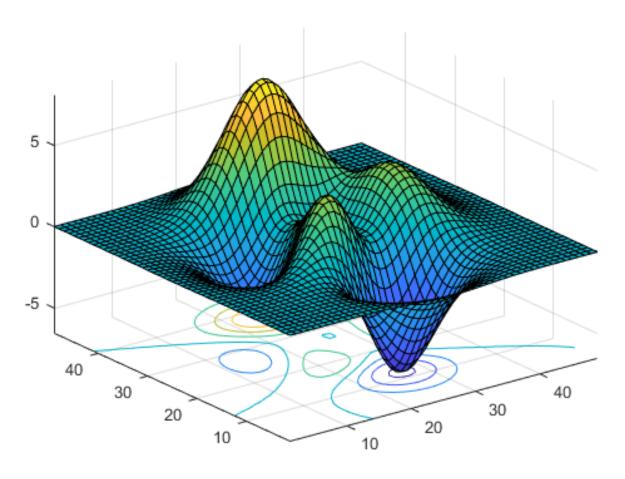
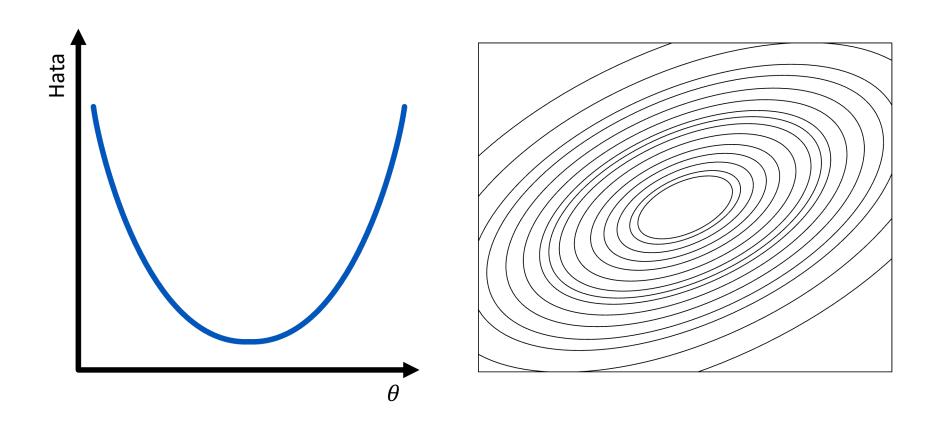
Contour (Eş yükselti eğrisi) Plot



https://www.mathworks.com/help/matlab/ref/surfc.html

Contour (Eş yükselti eğrisi) Plot



{Batch vs Mini-Batch vs Stochastic} Gradient Descent (BGD vs MGD vs SGD)

Cross-entropy Loss:

$$J(\theta) = -\frac{1}{m} \left[\sum_{i=1}^{m} y^{(i)} \log \hat{y}^{(i)} + (1 - y^{(i)}) \log(1 - \hat{y}^{(i)}) \right]$$

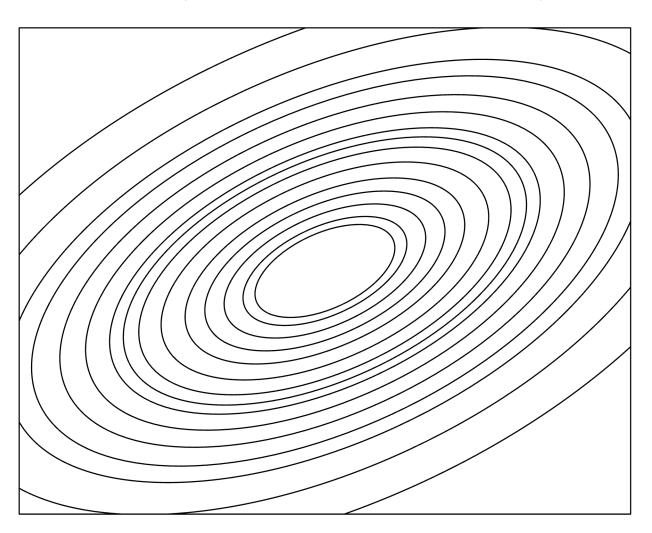
Gradient descent:

$$\frac{\partial}{\partial \theta_j} J(\theta) = \frac{1}{m} \sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)}) x_j^{(i)}$$

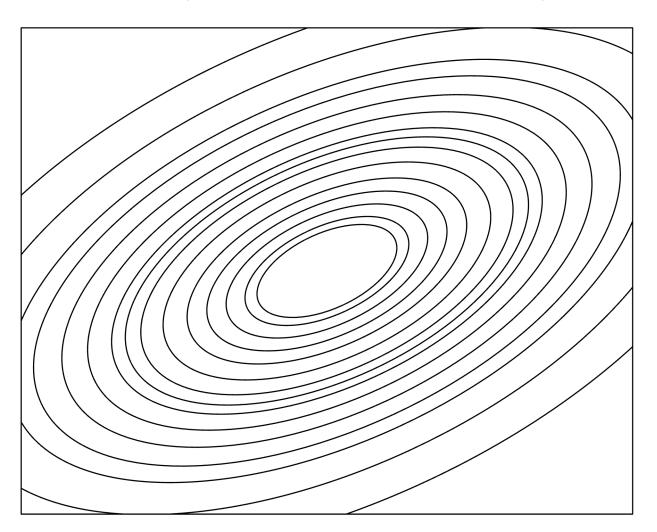
Yakınsayana kadar tekrar et {

$$\theta_j \coloneqq \theta_j - \alpha \frac{1}{m} \sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)}) x_j^{(i)}$$

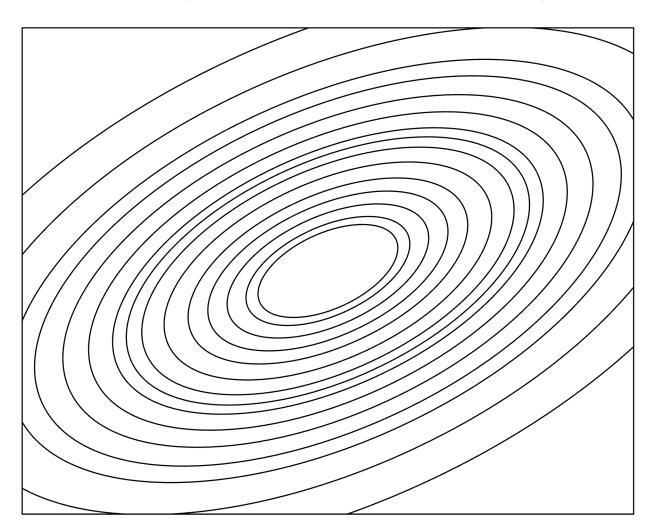
Batch vs Mini-Batch vs Stochastic Gradient Descent (BGD vs MGD vs SGD)



Batch vs Mini-Batch vs Stochastic Gradient Descent (BGD vs MGD vs SGD)



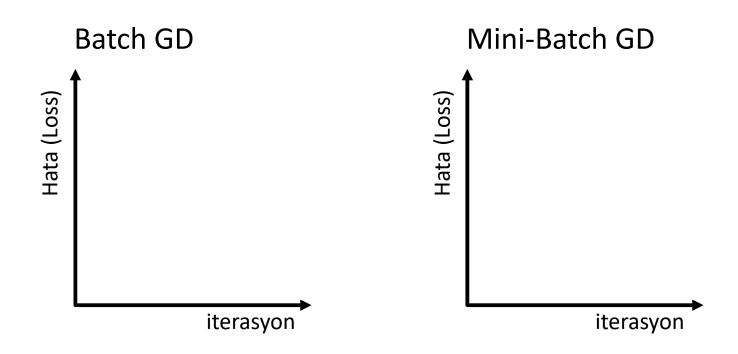
Batch vs Mini-Batch vs Stochastic Gradient Descent (BGD vs MGD vs SGD)



Batch Size

- Eğitim setinde 1500 adet 30 x 30 piksel resim olsun.
- batch size = 64 olsun
 - 1500 / 64 = 23 tam batch (her batch'te 64 resim) ve...
 - 1 adet 1500 (64 x 23) = 28 resimli batch olacaktır.
 - Toplamda 23 + 1 => 24 adet batch olacaktır.
- batch_size = 1500 iken ...
- batch_size = 1 iken ...

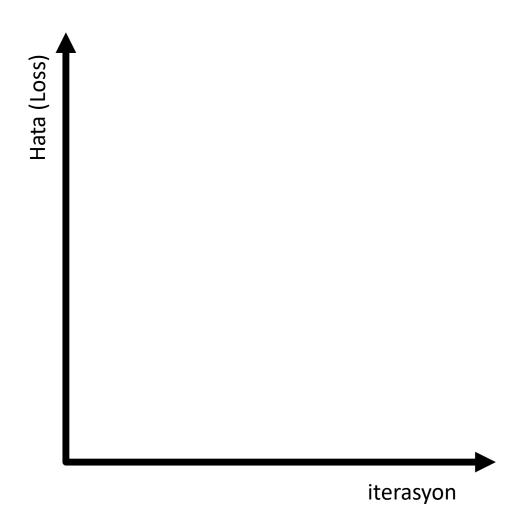
{Batch vs Mini-Batch vs Stochastic} Gradient Descent (BGD vs MGD vs SGD)



Epoch

- YSA'nın tüm 24 batch'e bakıp optimal çözüme doğru 1 adım atmasına 1 «epoch» denir.
- YSA tüm resimlere ne kadar çok bakarsa optimal çözüme o kadar yaklaşır.
 - Dikkat: Overfitting / Ezberleme durumu!!!
- Ezberlemenin (Overfitting / High variance) önüne geçebilmek için kullanılan yöntemlerden biri eğitim sırasında Validasyon setinin hata grafiğini kullanmaktır (Erken durma / early stopping).

Epoch



ML Mimarileri

https://playground.tensorflow.org

https://ml-playground.com