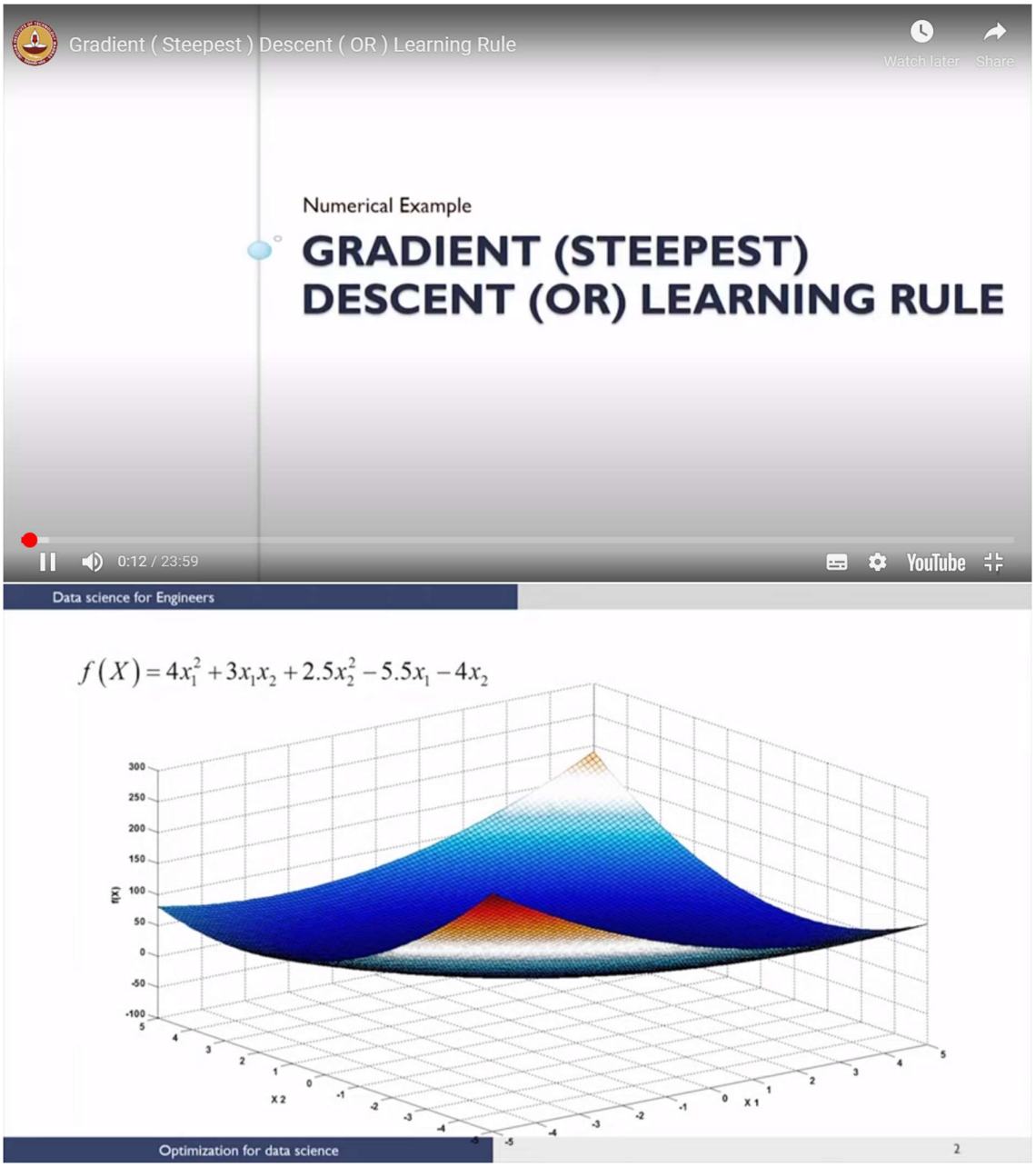

Gradient (Steepest) Descent (OR) Learning Rule
Watch later
Share
Numerical Example
GRADIENT (STEEPEST) DESCENT (OR) LEARNING RULE
0:12 / 23:59
YouTube
Data science for Engineers

$f(X) = 4x_1^2 + 3x_1x_2 + 2.5x_2^2 - 5.5x_1 - 4x_2$
f(X)
X 1
X 2
Optimization for data science
2

$$f'(X) = \begin{bmatrix} 8x_1 + 3x_2 - 5.5 \\ 3x_1 + 5x_2 - 4 \end{bmatrix}$$

Learning parameter $(\alpha) = 0.135$

Initial guess $(X_0) = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ $f(X_0) = 19$

Step 1: $X_1 = X_0 - \alpha f'(X_0)$

Gradient Descent (or) Learning Rule in ML

$$X_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix} - 0.135 \begin{bmatrix} 8x_{0,1} + 3x_{0,2} - 5.5 \\ 3x_{0,1} + 5x_{0,2} - 4 \end{bmatrix}$$

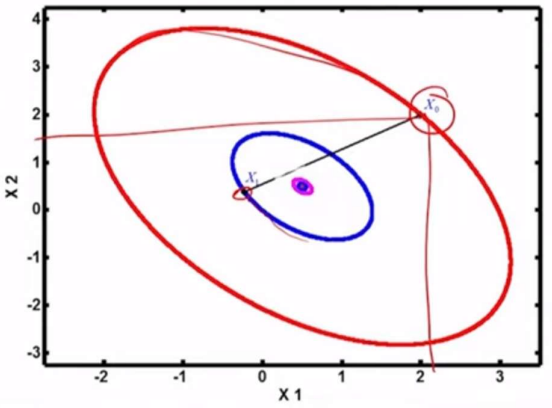
$$X_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix} - 0.135 \begin{bmatrix} 8(2) + 3(2) - 5.5 \\ 3(2) + 5(2) - 4 \end{bmatrix}$$

$$X_1 = \begin{bmatrix} -0.2275 \\ 0.3800 \end{bmatrix} \quad f(X_1) = 0.0399$$

Constant objective function contour plots

$$f(X) = 4x_1^2 + 3x_1x_2 + 2.5x_2^2 - 5.5x_1 - 4x_2 = K$$

Quadratic in this case - ellipse



Gradient (Steepest) Descent (OR) Learning Rule

First iteration $(X_1) = \begin{bmatrix} -0.2275 \\ 0.3800 \end{bmatrix}$

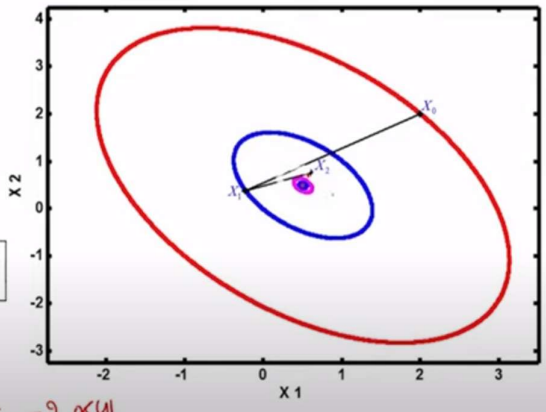
Step 2: $X_2 = X_1 - \alpha f'(X_1)$

$$X_2 = \begin{bmatrix} -0.2275 \\ 0.3800 \end{bmatrix} - 0.135 \begin{bmatrix} 8x_{1,1} + 3x_{1,2} - 5.5 \\ 3x_{1,1} + 5x_{1,2} - 4 \end{bmatrix}$$

$$X_2 = \begin{bmatrix} -0.2275 \\ 0.3800 \end{bmatrix} - 0.135 \begin{bmatrix} 8(-0.2275) + 3(0.3800) - 5.5 \\ 3(-0.2275) + 5(0.3800) - 4 \end{bmatrix}$$

$$X_2 = \begin{bmatrix} 0.6068 \\ 0.7556 \end{bmatrix} \quad f(X_2) = -2.0841$$

$f(x) = -2.0841$



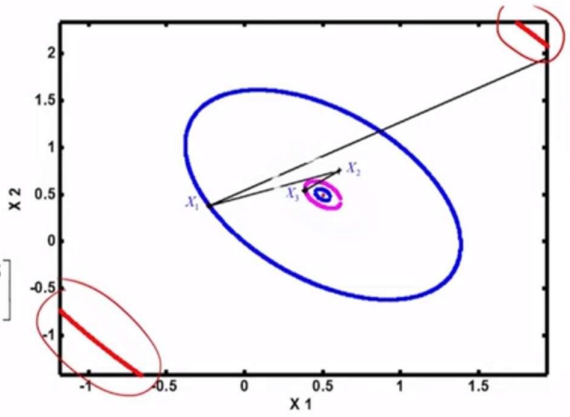
Second iteration $(X_2)=\begin{bmatrix}0.6068\\0.7556\end{bmatrix}$

Step 3: $X_3 = X_2 - \alpha f'(X_2)$

$$X_3 = \begin{bmatrix}0.6068\\0.7556\end{bmatrix} - 0.135\begin{bmatrix}8x_{2,1}+3x_{2,2}-5.5\\3x_{2,1}+5x_{2,2}-4\end{bmatrix}$$

$$X_3 = \begin{bmatrix}0.6068\\0.7556\end{bmatrix} - 0.135\begin{bmatrix}8(0.6068)+3(0.7556)-5.5\\3(0.6068)+5(0.7556)-4\end{bmatrix}$$

$$X_3 = \begin{bmatrix}0.3879\\0.5398\end{bmatrix} \qquad f(X_3) = -2.3342$$



Third iteration $(X_3)=\begin{bmatrix}0.3879\\0.5398\end{bmatrix}$

Step 4: $X_4 = X_3 - \alpha f'(X_3)$

$$X_4 = \begin{bmatrix}0.3879\\0.5398\end{bmatrix} - 0.135\begin{bmatrix}8x_{3,1}+3x_{3,2}-5.5\\3x_{3,1}+5x_{3,2}-4\end{bmatrix}$$

$$X_4 = \begin{bmatrix}0.3879\\0.5398\end{bmatrix} - 0.135\begin{bmatrix}8(0.3879)+3(0.5398)-5.5\\3(0.3879)+5(0.5398)-4\end{bmatrix}$$

$$X_4 = \begin{bmatrix}0.4928\\0.5583\end{bmatrix} \qquad f(X_4) = -2.3675$$

Optimal solution $(X_{opti})=\begin{bmatrix}0.5\\0.5\end{bmatrix}$ $f(X_{opti}) = -2.3750$

Gradient is zero at the optimum point