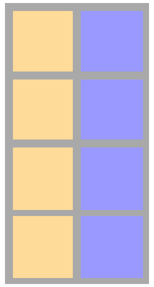
VEKTÖR ANALİZİ

2/2/2021, DERS 2

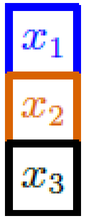
İlker Birbil

TENSÖRLER

$\boldsymbol{A} \in \mathbb{R}^{4\times 2}$



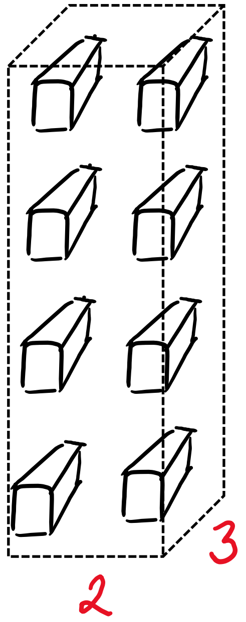
$\boldsymbol{x} \in \mathbb{R}^{3}$



$\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}\boldsymbol{x}} \in \mathbb{R}^{?}$

$$A = \begin{bmatrix} A_{11}(x) & A_{12}(x) \\ A_{21}(x) & A_{22}(x) \\ A_{31}(x) & A_{32}(x) \\ A_{41}(x) & A_{42}(x) \end{bmatrix}$$

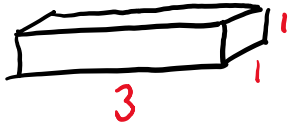
$$\frac{dA}{dx} = \begin{bmatrix} \nabla_x A_{11} & \nabla_x A_{12} \\ \nabla_x A_{21} & \nabla_x A_{22} \\ \nabla_x A_{31} & \nabla_x A_{32} \\ \nabla_x A_{41} & \nabla_x A_{42} \end{bmatrix} =$$



TENSÖR

$$A_{ij} : \mathbb{R}^3 \to \mathbb{R}$$

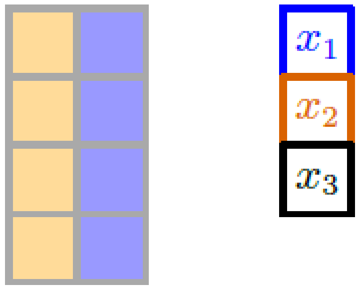
$$\nabla_x A_{ij} = \left[\frac{\partial A_{ij}}{\partial x_1} \quad \frac{\partial A_{ij}}{\partial x_2} \quad \frac{\partial A_{ij}}{\partial x_3} \right]$$



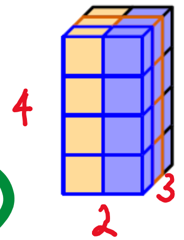
$$\frac{dA}{dx} \in \mathbb{R}^{4\times 2\times 3}$$

TENSÖRLER

$\boldsymbol{A} \in \mathbb{R}^{4\times 2}$ $\quad$ $\boldsymbol{x} \in \mathbb{R}^{3}$

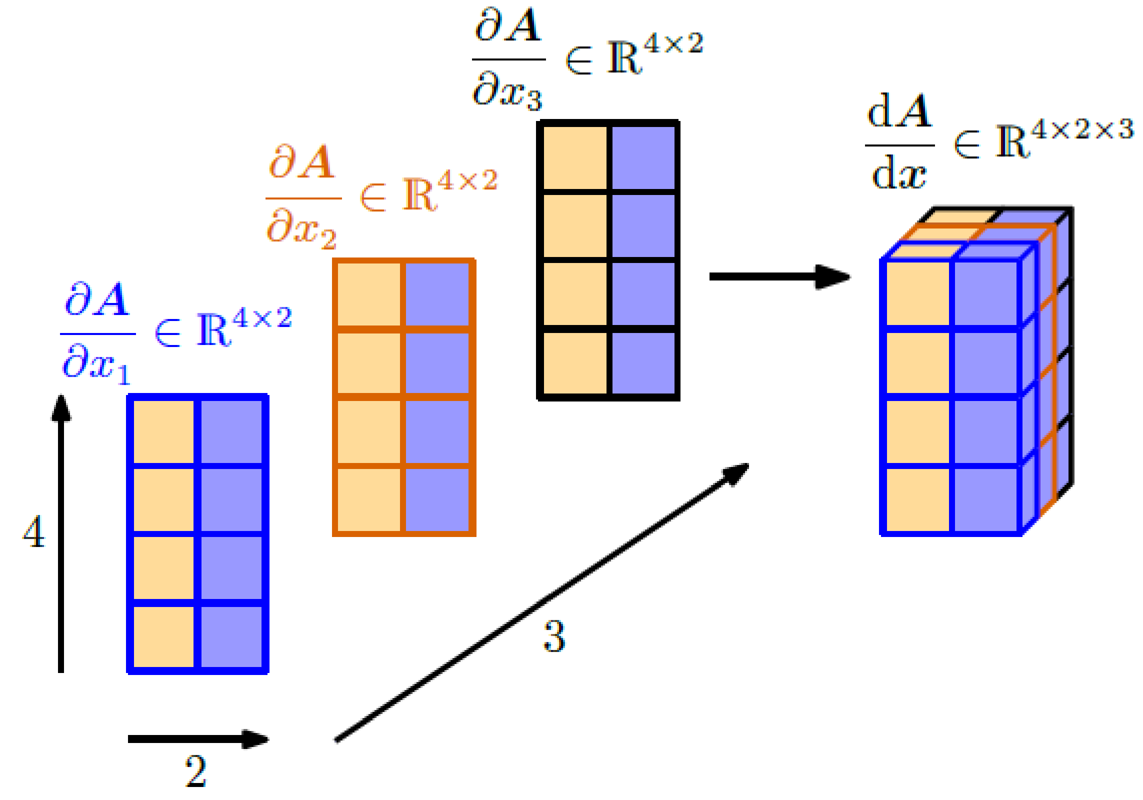


$$\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}\boldsymbol{x}} \in \mathbb{R}^{4\times 2\times 3}$$



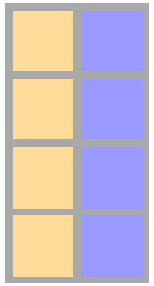
TENSÖR

1. YOL

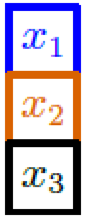


TENSÖRLER

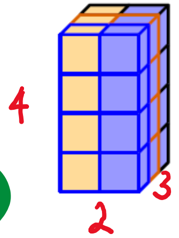
$\boldsymbol{A} \in \mathbb{R}^{4\times 2}$



$\boldsymbol{x} \in \mathbb{R}^{3}$



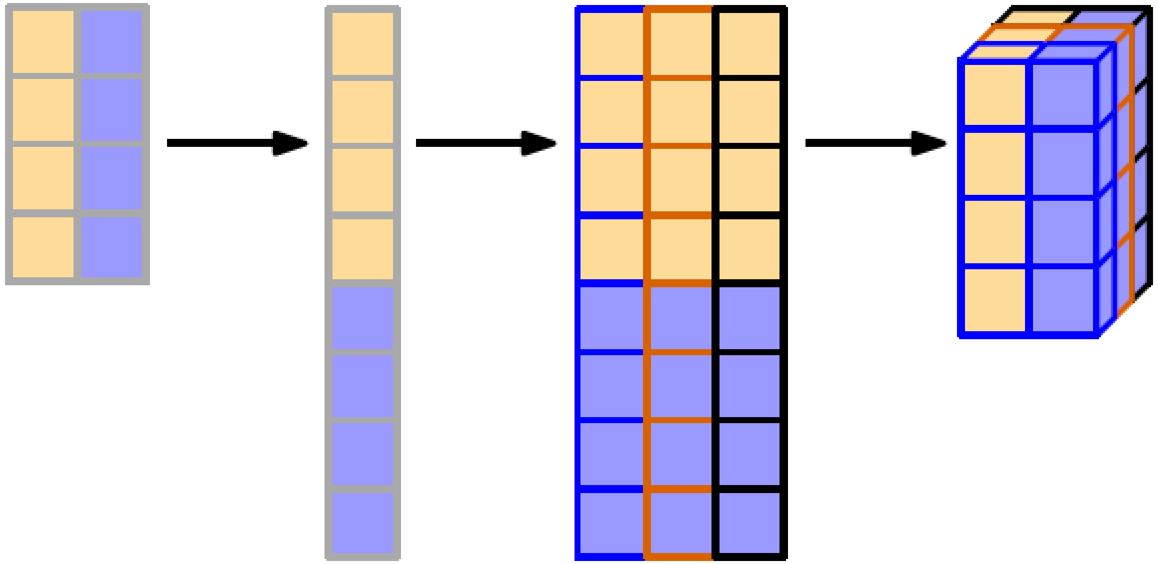
$$\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}\boldsymbol{x}} \in \mathbb{R}^{4\times 2\times 3}$$



TENSÖR

2. YOL

$\boldsymbol{A} \in \mathbb{R}^{4\times 2}$ $\qquad \tilde{\boldsymbol{A}} \in \mathbb{R}^{8}$ $\qquad \dfrac{\mathrm{d}\tilde{\boldsymbol{A}}}{\mathrm{d}\boldsymbol{x}} \in \mathbb{R}^{8\times 3}$ $\qquad \dfrac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}\boldsymbol{x}} \in \mathbb{R}^{4\times 2\times 3}$



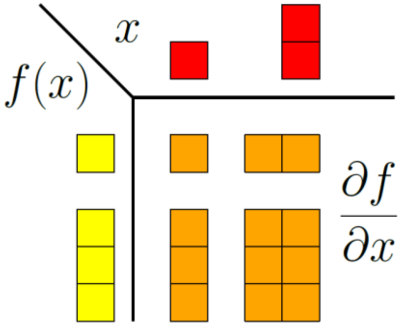
TENSÖRLER

$$\boldsymbol{f} = \boldsymbol{A}\boldsymbol{x}\,, \quad \boldsymbol{f}\in\mathbb{R}^{M}, \quad \boldsymbol{A}\in\mathbb{R}^{M\times N}, \quad \boldsymbol{x}\in\mathbb{R}^{N}$$

$$\frac{\mathrm{d}\boldsymbol{f}}{\mathrm{d}\boldsymbol{A}}\in\mathbb{R}^{?}$$

$$f:\mathbb{R}^{M\times N}\to\mathbb{R}^{M}\,, \qquad \frac{df}{\partial A} = \begin{array}{cc} & M\times N \\ M & \begin{bmatrix} \quad & \quad \end{bmatrix} \end{array} \in \mathbb{R}^{M\times(M\times N)}$$

TENSÖR



$$\boldsymbol{f}(\boldsymbol{R}) = \boldsymbol{R}^{\top}\boldsymbol{R} =: \boldsymbol{K}\in\mathbb{R}^{N\times N}\,, \quad \boldsymbol{R}\in\mathbb{R}^{M\times N}$$

$$\frac{\mathrm{d}\boldsymbol{K}}{\mathrm{d}\boldsymbol{R}}\in\mathbb{R}^{?}$$

$$f:\mathbb{R}^{M\times N}\to\mathbb{R}^{N\times N}\,, \qquad \frac{dK}{dR} = \begin{array}{cc} & M\times N \\ N\times N & \begin{bmatrix} \quad & \quad \end{bmatrix} \end{array} \in \mathbb{R}^{(N\times N)\times(M\times N)}$$

TENSÖR

KISA YOLLAR

$$\frac{\partial}{\partial \boldsymbol{X}} \boldsymbol{f}(\boldsymbol{X})^{\top} = \left(\frac{\partial \boldsymbol{f}(\boldsymbol{X})}{\partial \boldsymbol{X}}\right)^{\top}$$

$$\frac{\partial}{\partial \boldsymbol{X}} \operatorname{tr}(\boldsymbol{f}(\boldsymbol{X})) = \operatorname{tr}\left(\frac{\partial \boldsymbol{f}(\boldsymbol{X})}{\partial \boldsymbol{X}}\right)$$

$$\frac{\partial}{\partial \boldsymbol{X}} \det(\boldsymbol{f}(\boldsymbol{X})) = \det(\boldsymbol{f}(\boldsymbol{X})) \operatorname{tr}\left(\boldsymbol{f}(\boldsymbol{X})^{-1} \frac{\partial \boldsymbol{f}(\boldsymbol{X})}{\partial \boldsymbol{X}}\right)$$

$$\frac{\partial}{\partial \boldsymbol{X}} \boldsymbol{f}(\boldsymbol{X})^{-1} = -\boldsymbol{f}(\boldsymbol{X})^{-1} \frac{\partial \boldsymbol{f}(\boldsymbol{X})}{\partial \boldsymbol{X}} \boldsymbol{f}(\boldsymbol{X})^{-1}$$

$$\frac{\partial \boldsymbol{a}^{\top} \boldsymbol{X}^{-1} \boldsymbol{b}}{\partial \boldsymbol{X}} = -(\boldsymbol{X}^{-1})^{\top} \boldsymbol{a} \boldsymbol{b}^{\top} (\boldsymbol{X}^{-1})^{\top}$$

$$\frac{\partial \boldsymbol{x}^{\top} \boldsymbol{a}}{\partial \boldsymbol{x}} = \boldsymbol{a}^{\top}$$

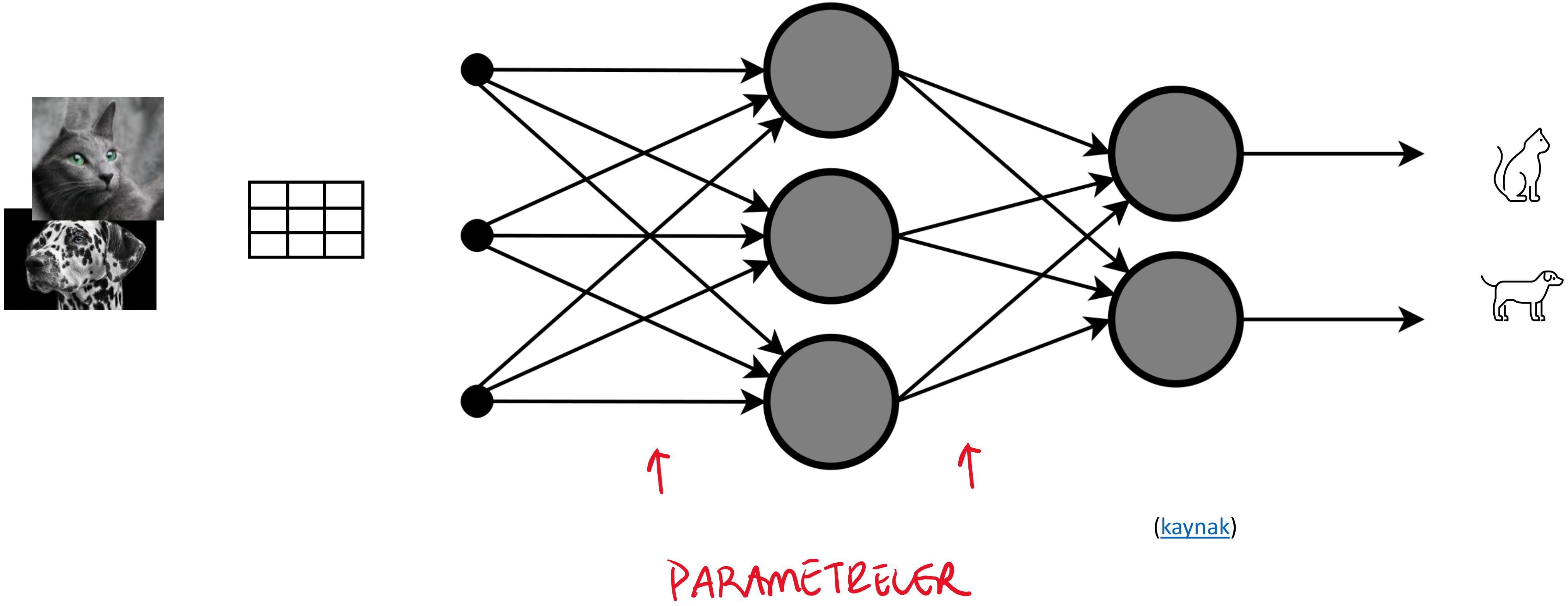
$$\frac{\partial \boldsymbol{a}^{\top} \boldsymbol{x}}{\partial \boldsymbol{x}} = \boldsymbol{a}^{\top}$$

$$\frac{\partial \boldsymbol{a}^{\top} \boldsymbol{X} \boldsymbol{b}}{\partial \boldsymbol{X}} = \boldsymbol{a} \boldsymbol{b}^{\top}$$

$$\frac{\partial \boldsymbol{x}^{\top} \boldsymbol{B} \boldsymbol{x}}{\partial \boldsymbol{x}} = \boldsymbol{x}^{\top} (\boldsymbol{B} + \boldsymbol{B}^{\top})$$

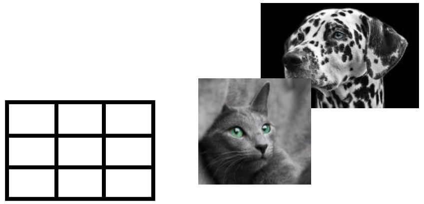
$$\frac{\partial}{\partial \boldsymbol{s}} (\boldsymbol{x} - \boldsymbol{A}\boldsymbol{s})^{\top} \boldsymbol{W} (\boldsymbol{x} - \boldsymbol{A}\boldsymbol{s}) = -2(\boldsymbol{x} - \boldsymbol{A}\boldsymbol{s})^{\top} \boldsymbol{W} \boldsymbol{A}$$ $(W^T = W)$

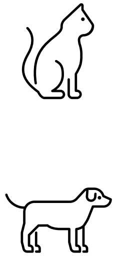
GERİ YAYILIM



(kaynak)

GERİ YAYILIM



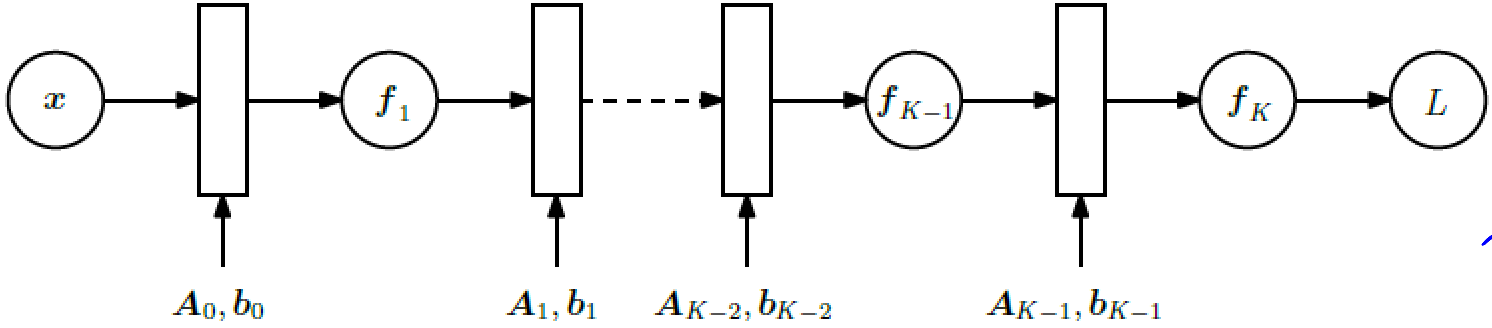


$$\left.\begin{matrix} 0 \\ 1 \end{matrix}\right\} y$$

$$(f_K \circ f_{K-1} \circ \cdots \circ f_1)(\boldsymbol{x}) = f_K(f_{K-1}(\cdots(f_1(\boldsymbol{x}))\cdots))$$

$$\boldsymbol{f}_0 := \boldsymbol{x}$$

$$\boldsymbol{f}_i := \sigma_i(\boldsymbol{A}_{i-1}\boldsymbol{f}_{i-1} + \boldsymbol{b}_{i-1}), \quad i = 1, \ldots, K$$



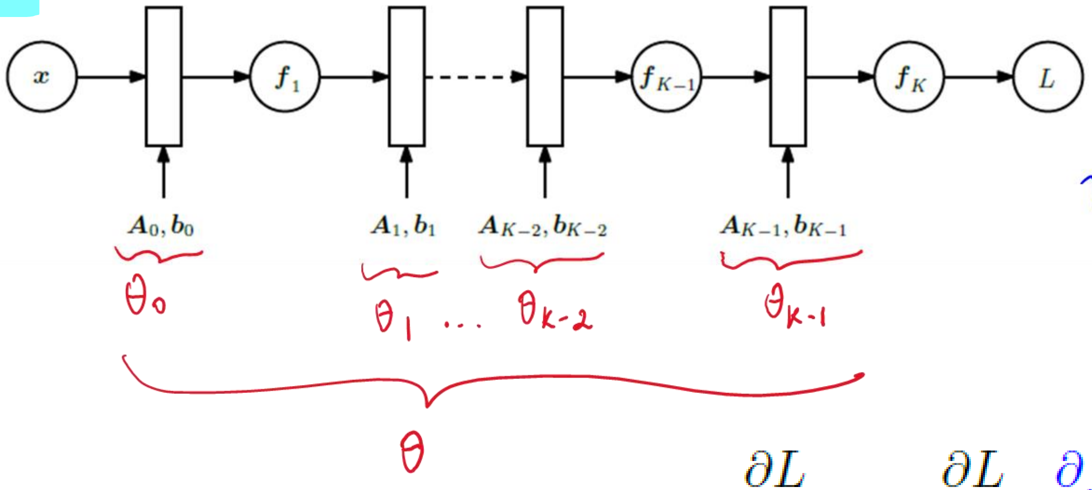
θ_0 $\quad\theta_1 \ldots \theta_{K-2}$ $\quad\theta_{K-1}$

θ PARAMETRELER

$$\min_{\theta} \{ \underbrace{\|y - f_K(\theta)\|^2}_{L(\theta)} \}$$

$$\frac{\partial L}{\partial \theta} = ?$$

GERİ YAYILIM

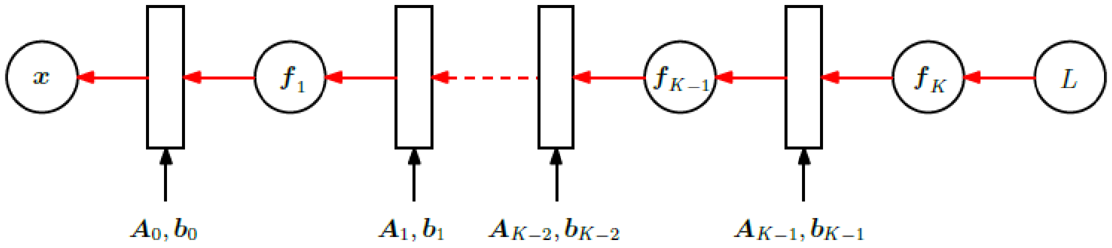


$$\min_{\theta} \{ \underbrace{\|y - f_K(\theta)\|^2}_{L(\theta)} \}$$

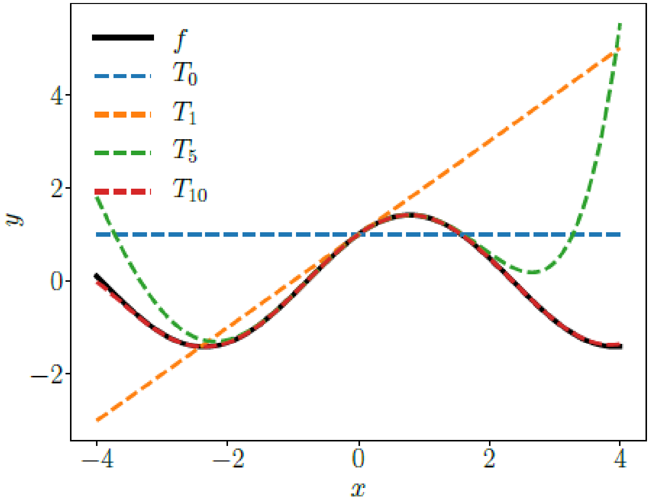
$$\frac{\partial L}{\partial \boldsymbol{\theta}_{K-1}} = \frac{\partial L}{\partial \boldsymbol{f}_K} \frac{\partial \boldsymbol{f}_K}{\partial \boldsymbol{\theta}_{K-1}}$$

$$\frac{\partial L}{\partial \boldsymbol{\theta}_{K-2}} = \frac{\partial L}{\partial \boldsymbol{f}_K} \boxed{\frac{\partial \boldsymbol{f}_K}{\partial \boldsymbol{f}_{K-1}} \frac{\partial \boldsymbol{f}_{K-1}}{\partial \boldsymbol{\theta}_{K-2}}}$$

$$\frac{\partial L}{\partial \boldsymbol{\theta}_{K-3}} = \frac{\partial L}{\partial \boldsymbol{f}_K} \frac{\partial \boldsymbol{f}_K}{\partial \boldsymbol{f}_{K-1}} \boxed{\frac{\partial \boldsymbol{f}_{K-1}}{\partial \boldsymbol{f}_{K-2}} \frac{\partial \boldsymbol{f}_{K-2}}{\partial \boldsymbol{\theta}_{K-3}}}$$



İKİNCİ DERECE



$$f : \mathbb{R}^n \to \mathbb{R} \qquad \nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix}$$

$$\nabla^2 f = H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & & \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix} \qquad (H = H^T)$$

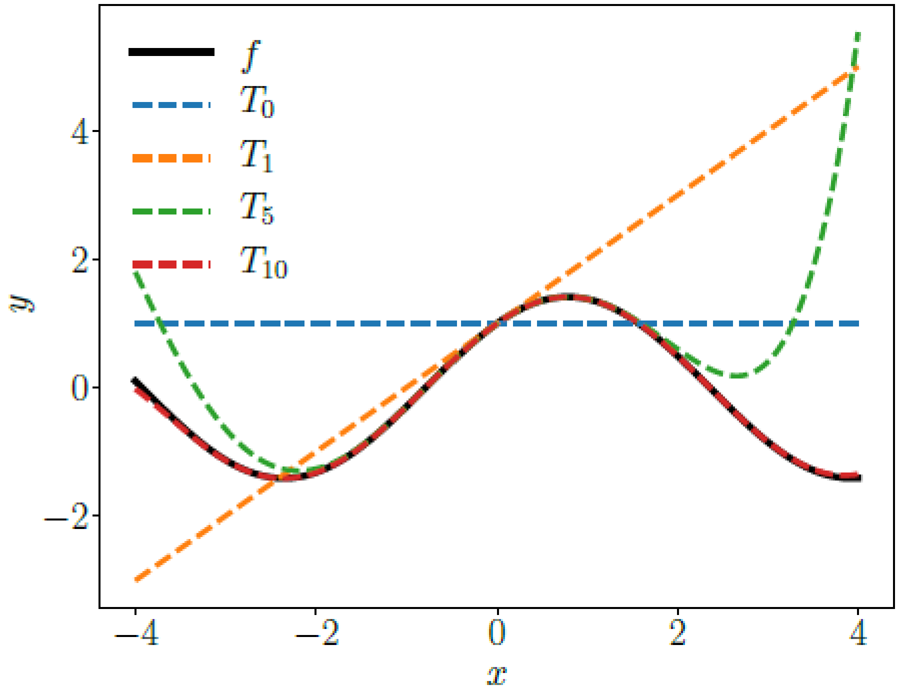
HESYEN MATRİSİ

$$f(x) \approx f(x_0) + \nabla f(x_0)(x - x_0) + \frac{1}{2}(x - x_0)^T \nabla^2 f(x_0)(x - x_0)$$

İKİNCİ DERECE

$x \in \mathbb{R}^n$

$$f(x) \approx f(x_0) + \nabla f(x_0)(x - x_0) + \frac{1}{2}(x - x_0)^T \nabla^2 f(x_0)(x - x_0)$$



$$T_\infty(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k.$$

$$f(x) \approx f(x_0) + f^{(1)}(x_0)(x - x_0) + \underbrace{\frac{1}{2} f^{(2)}(x_0)(x - x_0)^2}_{\frac{1}{2}(x - x_0) f^{(2)}(x_0)(x - x_0)}$$

$x \in \mathbb{R}$

ÖDEV

5.4 Compute the Taylor polynomials T_n, $n = 0, \ldots, 5$ of $f(x) = \sin(x) + \cos(x)$

5.6 Differentiate f with respect to $\boldsymbol{t}$ and g with respect to $\boldsymbol{X}$, where

5.9 We define

$$g(\boldsymbol{z}, \boldsymbol{\nu}) := \log p(\boldsymbol{x}, \boldsymbol{z}) - \log q(\boldsymbol{z}, \boldsymbol{\nu})$$
$$\boldsymbol{z} := t(\boldsymbol{\epsilon}, \boldsymbol{\nu})$$

for differentiable functions p, q, t, and $\boldsymbol{x} \in \mathbb{R}^D, \boldsymbol{z} \in \mathbb{R}^E, \boldsymbol{\nu} \in \mathbb{R}^F, \boldsymbol{\epsilon} \in \mathbb{R}^G$. By using the chain rule, compute the gradient

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\nu}} g(\boldsymbol{z}, \boldsymbol{\nu}) \,.$$