Introduction to Tensor

CE40282-1: Linear Algebra
Hamid R. Rabiee and Maryam Ramezani
Sharif University of Technology

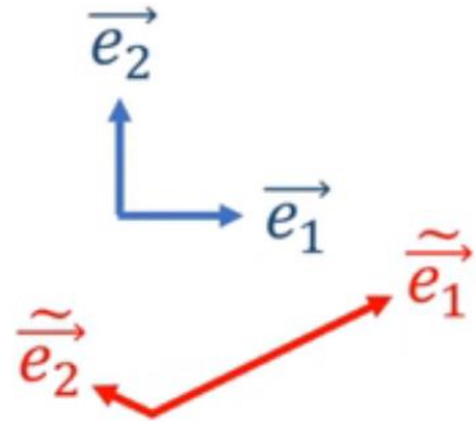
Covector

- Co~~n~~vectors (covariant vectors) are basically row vectors in an orthonormal basis, but not in general!

ستون

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 & 1 \end{bmatrix}$$

row سطر



- Row vectors are function on column vectors.

$$\alpha : V \to \mathbb{R}$$

$$\alpha(\vec{v}) = \alpha_1 v^1 + \alpha_2 v^2 + \cdots + \alpha_n v^n$$

$$\sum_{i=1}^{n} \alpha_i v^i$$

Covector

- A function that takes a vector and produces a scalar.

- Properties:

خاصیت خطی:

$$\alpha(\vec{v} + \vec{w}) = \alpha(\vec{v}) + \alpha(\vec{w})$$

$$\alpha(n\vec{v}) = n\alpha(\vec{v})$$

$$\alpha(n\vec{v} + m\vec{w}) = n\alpha(\vec{v}) + m\alpha(\vec{w})$$

Covector Visualization

$\alpha(v)$, 1 $\quad$ $2+2=4$

$\alpha = [2 \quad 1]\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = 2x + 1y$

$2x$

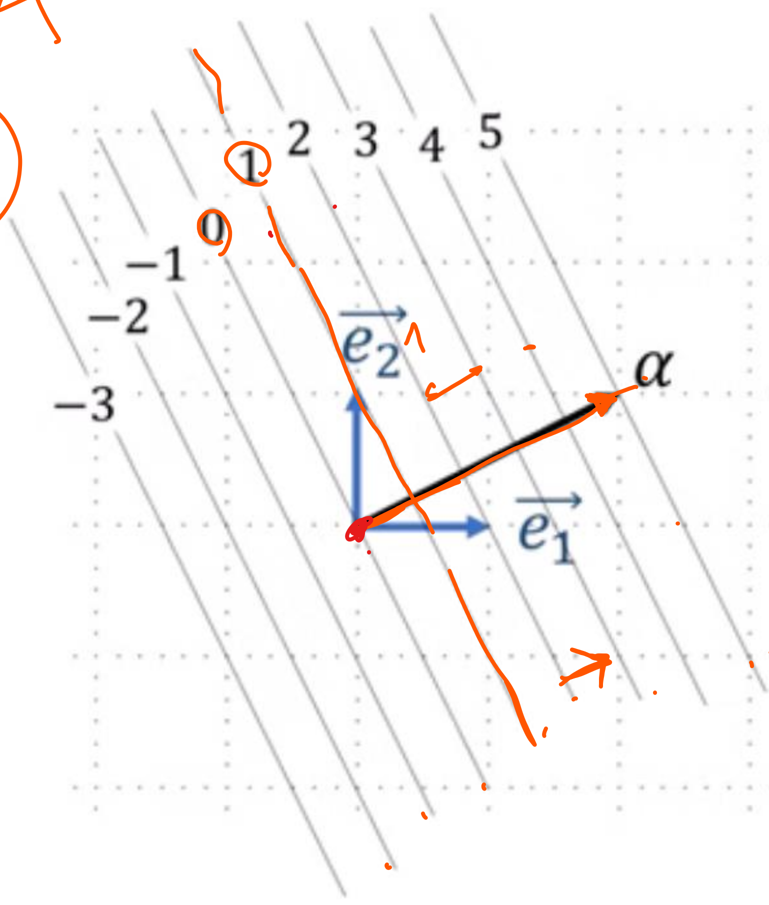
$2x + 1y = 0 \rightarrow y = -2x$

$2x + 1y = 1 \rightarrow y = 1 - 2x$

$2x + 1y = 2$

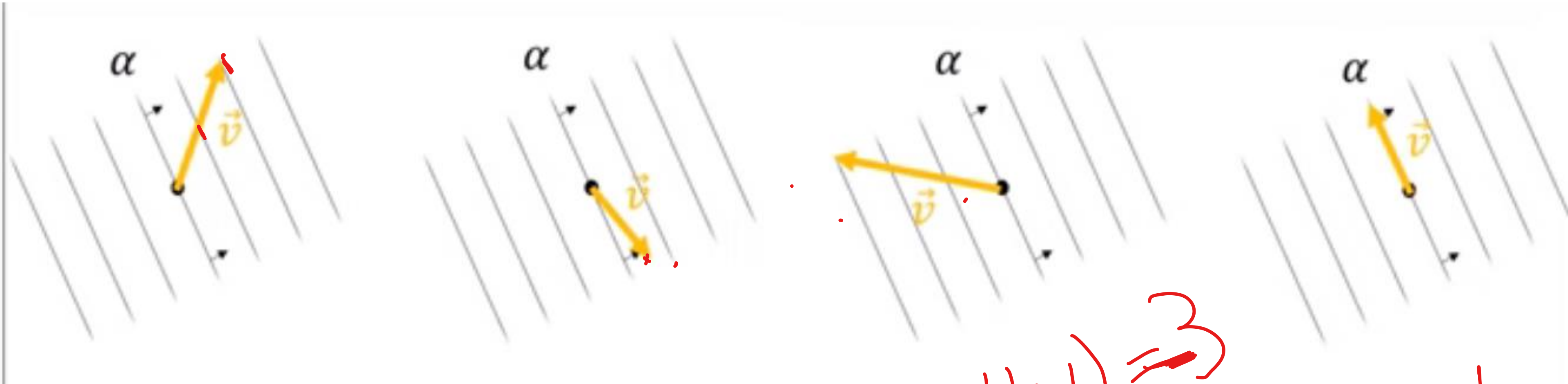
$2x + 1y = -1$

$2x + 1y = -2$



Covector Visualization

- Example of vector and convector



$\alpha(v)=2$ $\alpha(v)=0.5$ $\alpha(v)=3$ $\alpha(v)=0$

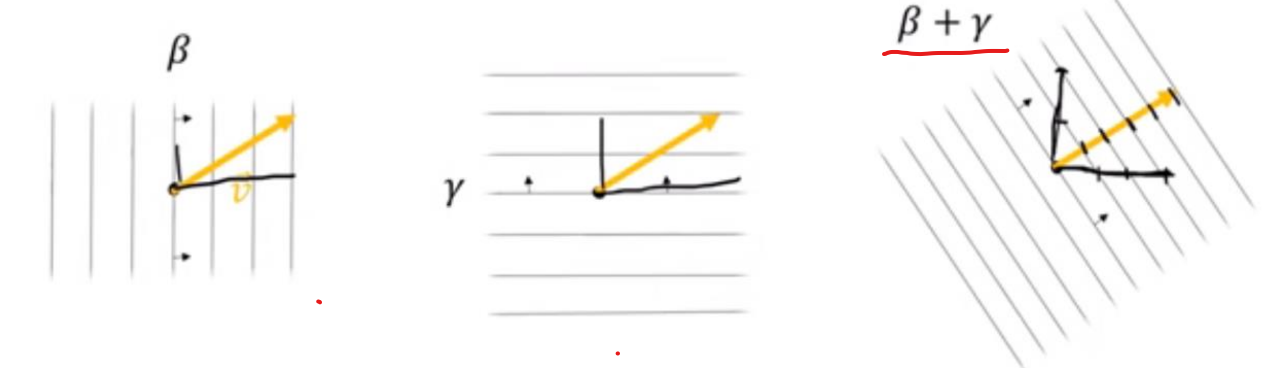
Covector Operations

$\alpha(2v)$

- Scaling

$\alpha(v)=2$ $2\alpha(v)=4$

- Adding



$\beta(\vec{v}) = 3$ $\gamma(\vec{v}) = 2$ $(\beta + \gamma)(\vec{v}) = 5$

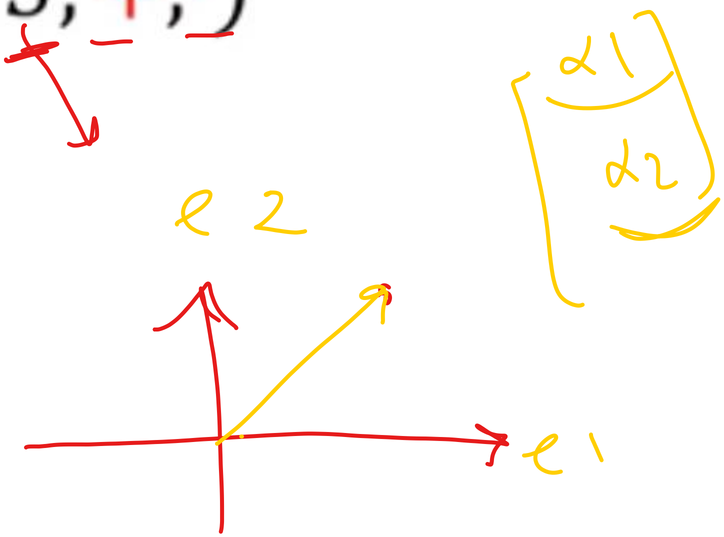
$(\beta + \gamma)(\vec{v}) = \beta(\vec{v}) + \gamma(\vec{v})$

Dual vector space

- Vector Space $(V, S, +, \cdot)$

  Dual Vector Space $(V^*, S, +, \cdot)$

covector



Covector Components

When we write $\begin{bmatrix} 2 \\ 1 \end{bmatrix}_{\vec{e_i}}$, we mean $2\vec{e_1} + 1\vec{e_2}$

What exactly are we measure with when we write $[2 \quad 1]$?

- Covectors are functions $\alpha : V \to \mathbb{R}$
- Covectors don't live in the vector space V
- We use can't basis vectors in V like $\{\vec{e_1}, \vec{e_2}\}$ to measure covectors

Take the basis $\{\vec{e_1}, \vec{e_2}\}$ for V.

kronecker delta

Introduce two special covectors $\epsilon^1, \epsilon^2 : V \to \mathbb{R}$

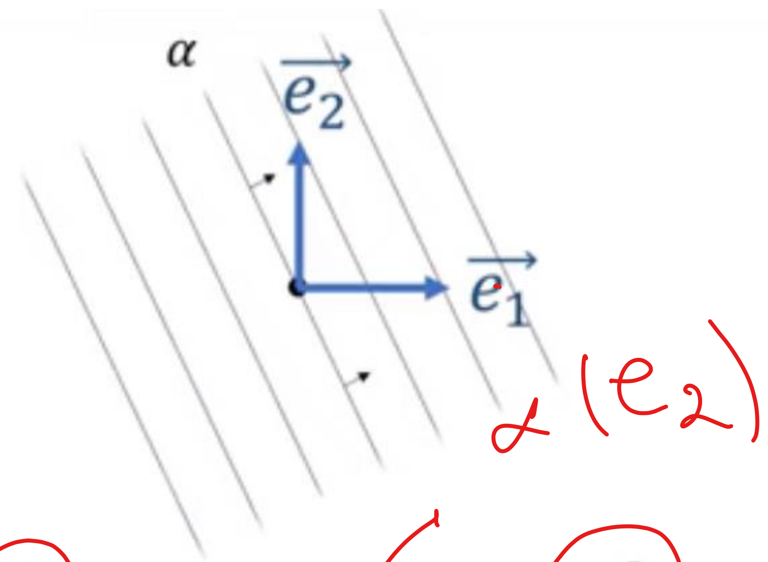
- $\epsilon^1(\vec{e_1}) = 1 \qquad \epsilon^1(\vec{e_2}) = 0$
- $\epsilon^2(\vec{e_1}) = 0 \qquad \epsilon^2(\vec{e_2}) = 1$

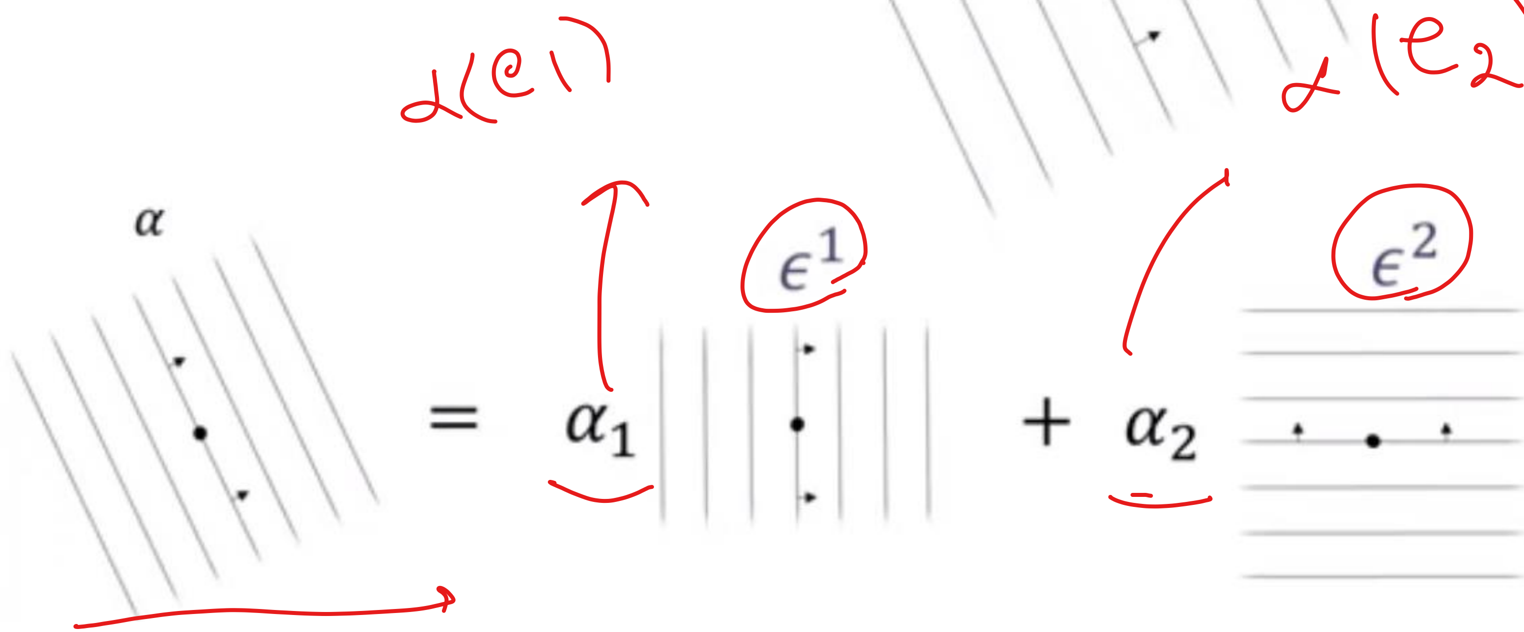
$$\epsilon^i(\vec{e_j}) = \delta_{ij} = \begin{cases} 1 \; if \; i = j \\ 0 \; if \; i \neq j \end{cases}$$

Dual Basis

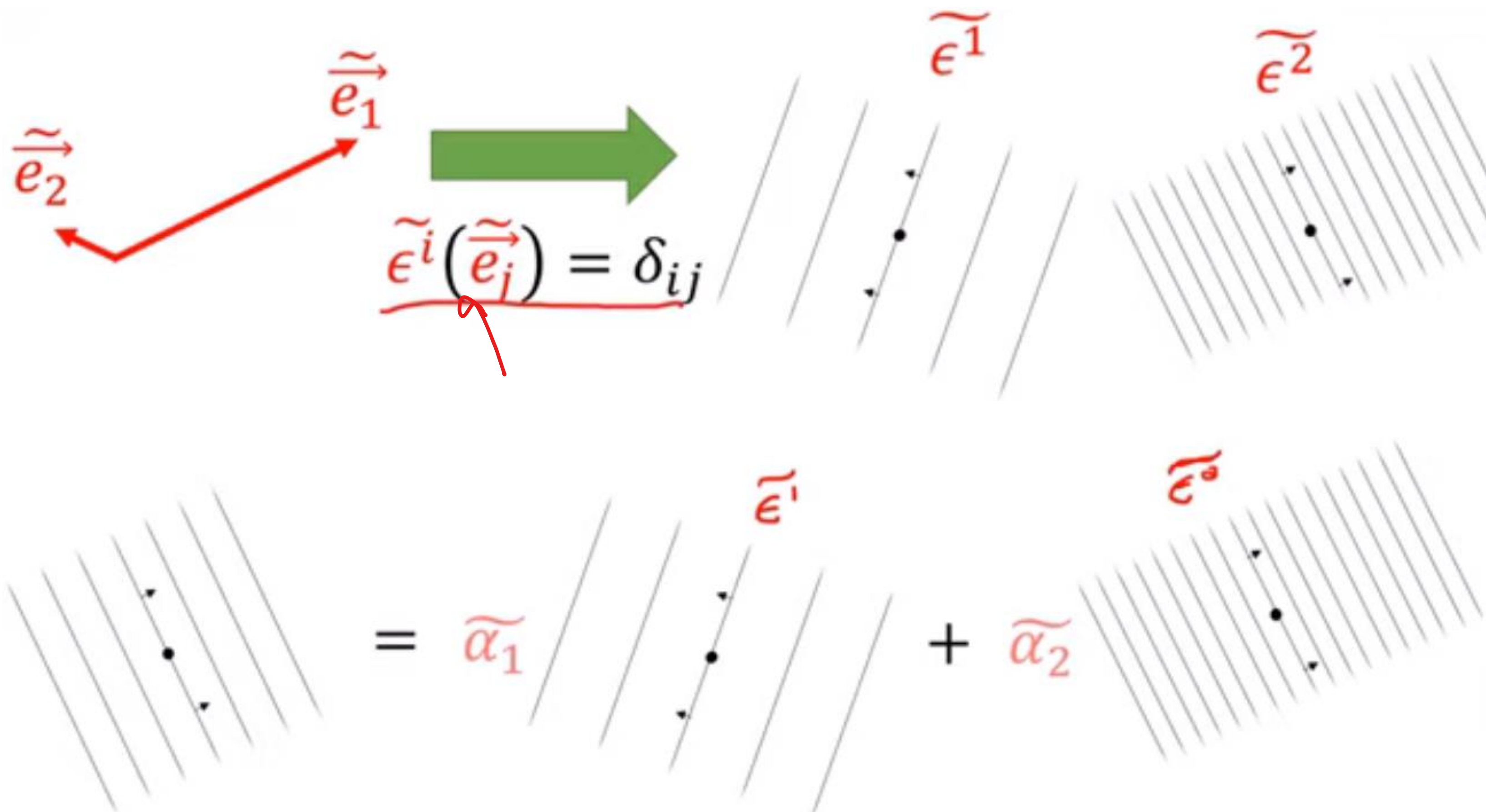
$$\alpha = \alpha_1 \epsilon^1 + \alpha_2 \epsilon^2$$

- $\alpha(\overrightarrow{e_1}) = \alpha_1$
- $\alpha(\overrightarrow{e_2}) = \alpha_2$





Dual Basis

- 
$\vec{\widetilde{e}}_1$
$\vec{\widetilde{e}}_2$
$\widetilde{\epsilon}^i(\vec{\widetilde{e}}_j) = \delta_{ij}$
$\widetilde{\epsilon}^1$
$\widetilde{\epsilon}^2$
$= \widetilde{\alpha}_1$
$\widetilde{\epsilon}^1$
$+ \widetilde{\alpha}_2$
$\widetilde{\epsilon}^2$

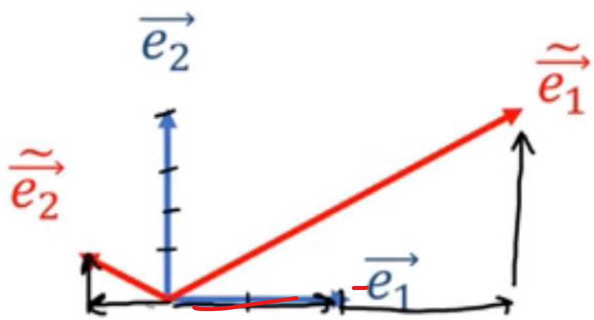
Coordination change

- $\alpha = 2\epsilon^1 + 1\epsilon^2$ $\qquad$ $\alpha = \widetilde{\alpha_1}\widetilde{\epsilon^1} + \widetilde{\alpha_2}\widetilde{\epsilon^2}$

  $[2 \quad 1]_{\epsilon^i}$ $\qquad$ $[\widetilde{\alpha_1} \quad \widetilde{\alpha_2}]_{\widetilde{\epsilon^i}}$

  - $\alpha(\vec{e_1}) = \alpha_1 = 2$
  - $\alpha(\vec{e_2}) = \alpha_2 = 1$

  - $\alpha(\widetilde{\vec{e_1}}) = \widetilde{\alpha_1}$
  - $\alpha(\widetilde{\vec{e_2}}) = \widetilde{\alpha_2}$



$$\begin{bmatrix} 1/4 & 1/2 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix}_{\epsilon^i} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{\widetilde{\epsilon^i}}$$

$$\begin{aligned} \widetilde{\alpha_1} &= \alpha(\widetilde{\vec{e_1}}) \\ &= \alpha(2\vec{e_1} + 1\vec{e_2}) \\ &= 2\alpha(\vec{e_1}) + 1\alpha(\vec{e_2}) \\ &= 2(2) + 1(1) \\ &= 5 \end{aligned}$$

$$\begin{aligned} \widetilde{\alpha_2} &= \alpha(\widetilde{\vec{e_2}}) \\ &= \alpha(-1/2\vec{e_1} + 1/4\vec{e_2}) \\ &= -1/2\alpha(\vec{e_1}) + 1/4\alpha(\vec{e_2}) \\ &= -1/2(2) + 1/4(1) \\ &= -3/4 \end{aligned}$$

$[2 \quad 1]_{\epsilon^i}$

$[5 \quad -3/4]_{\widetilde{\epsilon^i}}$

$$B \begin{bmatrix} 2 \\ 1 \end{bmatrix}_{\epsilon^i} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{\widetilde{\epsilon^i}}$$

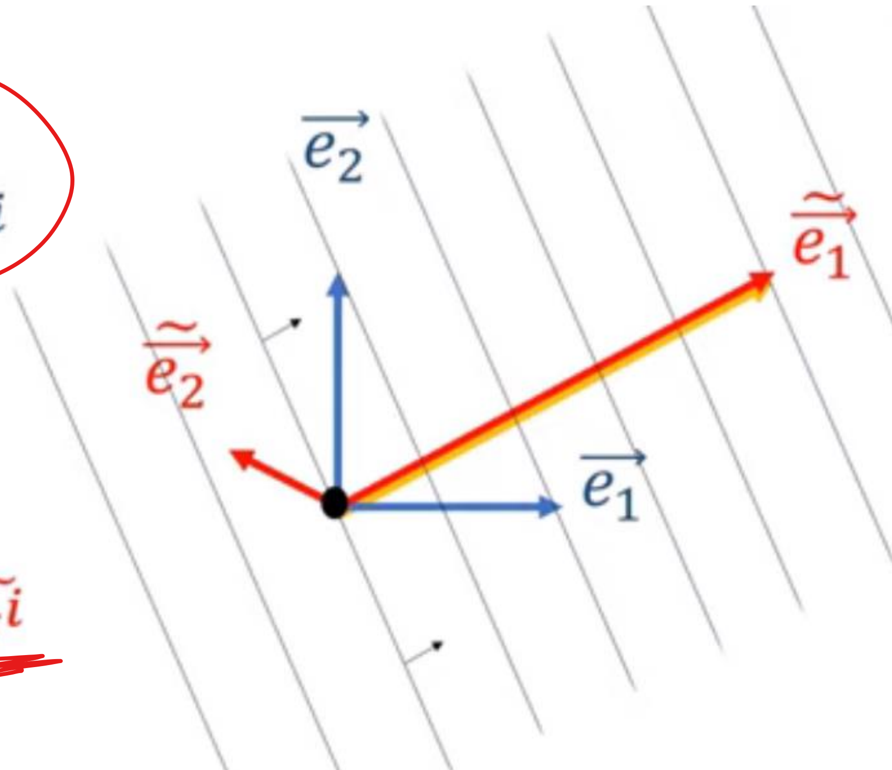
$$F \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{\widetilde{\epsilon^i}} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}_{\epsilon^i}$$

Conclusion

- Why "co~~n~~vector is a row vector" is not true?

$$v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}_{\epsilon^i} \longleftrightarrow \overset{\alpha(v)}{\begin{bmatrix} 2 & 1 \end{bmatrix}_{\epsilon^i}}$$

$$v = \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{\tilde{\epsilon}^i} \qquad \overset{\alpha(v)}{\underline{\begin{bmatrix} 5 & -3/4 \end{bmatrix}_{\tilde{\epsilon}^i}}}$$



Kronecker Product

Definition 14.1 *Let* $\boldsymbol{A}$ *be an* $m \times n$ *matrix and* $\boldsymbol{B}$ *a* $p \times q$ *matrix. The* **Kronecker product** *of* $\boldsymbol{A}$ *and* $\boldsymbol{B}$*, written as* $\boldsymbol{A} \otimes \boldsymbol{B}$*, is the* $mp \times nq$ *partitioned matrix whose* (i, j)*-th block is the* $p \times q$ *matrix* $a_{ij}\boldsymbol{B}$ *for* $i = 1, 2, \ldots, m$ *and* $j = 1, 2, \ldots, n$*. Thus,*

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{bmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \ldots & a_{1n}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & a_{22}\boldsymbol{B} & \ldots & a_{2n}\boldsymbol{B} \\ \vdots & \vdots & & \vdots \\ a_{m1}\boldsymbol{B} & a_{m2}\boldsymbol{B} & \ldots & a_{mn}\boldsymbol{B} \end{bmatrix} .$$

- Example

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 5 & 6 & 0 \\ 7 & 8 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 6 & 0 & 10 & 12 & 0 \\ 7 & 8 & 0 & 14 & 16 & 0 \\ 0 & 0 & 1 & 0 & 0 & 2 \\ 15 & 18 & 0 & 20 & 24 & 0 \\ 21 & 24 & 0 & 28 & 32 & 0 \\ 0 & 0 & 3 & 0 & 0 & 4 \\ 0 & 0 & 0 & 5 & 6 & 0 \\ 0 & 0 & 0 & 7 & 8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Kronecker Product

- Let $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ be any three matrices.

  (i) $\alpha \otimes \boldsymbol{A} = \boldsymbol{A} \otimes \alpha = \alpha \boldsymbol{A}$, for any scalar α.

  (ii) $(\alpha \boldsymbol{A}) \otimes (\beta \boldsymbol{B}) = \alpha\beta(\boldsymbol{A} \otimes \boldsymbol{B})$, for any scalars α and β.

  (iii) $(\boldsymbol{A} \otimes \boldsymbol{B}) \otimes \boldsymbol{C} = \boldsymbol{A} \otimes (\boldsymbol{B} \otimes \boldsymbol{C})$.

  (iv) $(\boldsymbol{A} + \boldsymbol{B}) \otimes \boldsymbol{C} = (\boldsymbol{A} \otimes \boldsymbol{C}) + (\boldsymbol{B} \otimes \boldsymbol{C})$, if $\boldsymbol{A}$ and $\boldsymbol{B}$ are of same size.

  (v) $\boldsymbol{A} \otimes (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{A} \otimes \boldsymbol{B}) + (\boldsymbol{A} \otimes \boldsymbol{C})$, if $\boldsymbol{B}$ and $\boldsymbol{C}$ are of same size.

  (vi) $(\boldsymbol{A} \otimes \boldsymbol{B})' = \boldsymbol{A}' \otimes \boldsymbol{B}'$.

- Other Properties:

$A \otimes B \neq B \otimes A$

$D \otimes A =$

$0 \otimes A = A \otimes 0 = 0$

- Ref: Chapter 14: Linear Algebra and Matrix Analysis for Statistics

Kronecker Product

$$\begin{bmatrix} v^1 \\ v^2 \end{bmatrix} \otimes \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix}$$

$$\begin{bmatrix} \begin{bmatrix} v^1 \\ v^2 \end{bmatrix} \alpha_1 & \begin{bmatrix} v^1 \\ v^2 \end{bmatrix} \alpha_2 \end{bmatrix}$$

$$\begin{bmatrix} \begin{bmatrix} v^1 \alpha_1 \\ v^2 \alpha_1 \end{bmatrix} & \begin{bmatrix} v^1 \alpha_2 \\ v^2 \alpha_2 \end{bmatrix} \end{bmatrix}$$

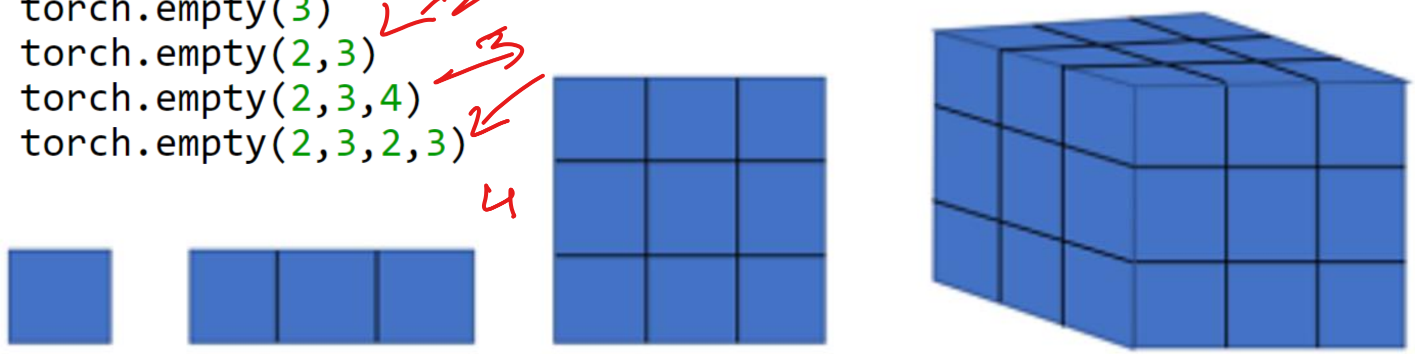
$$\begin{bmatrix} \begin{bmatrix} v^1 \alpha_1 \\ v^2 \alpha_1 \end{bmatrix} & \begin{bmatrix} v^1 \alpha_2 \\ v^2 \alpha_2 \end{bmatrix} \end{bmatrix} \otimes \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix}$$

$$\begin{bmatrix} \begin{bmatrix} \begin{bmatrix} v^1 \alpha_1 \\ v^2 \alpha_1 \end{bmatrix} & \begin{bmatrix} v^1 \alpha_2 \\ v^2 \alpha_2 \end{bmatrix} \end{bmatrix} \omega_1 \\ \begin{bmatrix} \begin{bmatrix} v^1 \alpha_1 \\ v^2 \alpha_1 \end{bmatrix} & \begin{bmatrix} v^1 \alpha_2 \\ v^2 \alpha_2 \end{bmatrix} \end{bmatrix} \omega_2 \end{bmatrix} \Rightarrow \begin{bmatrix} \begin{bmatrix} \begin{bmatrix} v^1 \alpha_1 \omega_1 \\ v^2 \alpha_1 \omega_1 \end{bmatrix} & \begin{bmatrix} v^1 \alpha_2 \omega_1 \\ v^2 \alpha_2 \omega_1 \end{bmatrix} \end{bmatrix} \\ \begin{bmatrix} \begin{bmatrix} v^1 \alpha_1 \omega_2 \\ v^2 \alpha_1 \omega_2 \end{bmatrix} & \begin{bmatrix} v^1 \alpha_2 \omega_2 \\ v^2 \alpha_2 \omega_2 \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Tensor

- Multi-dimensional array of numbers

```
w = torch.empty(3)
x = torch.empty(2,3)
y = torch.empty(2,3,4)
z = torch.empty(2,3,2,3)
```

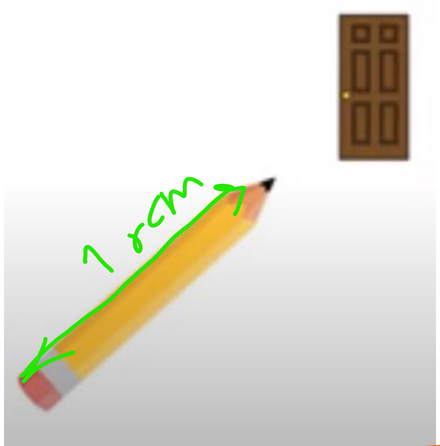


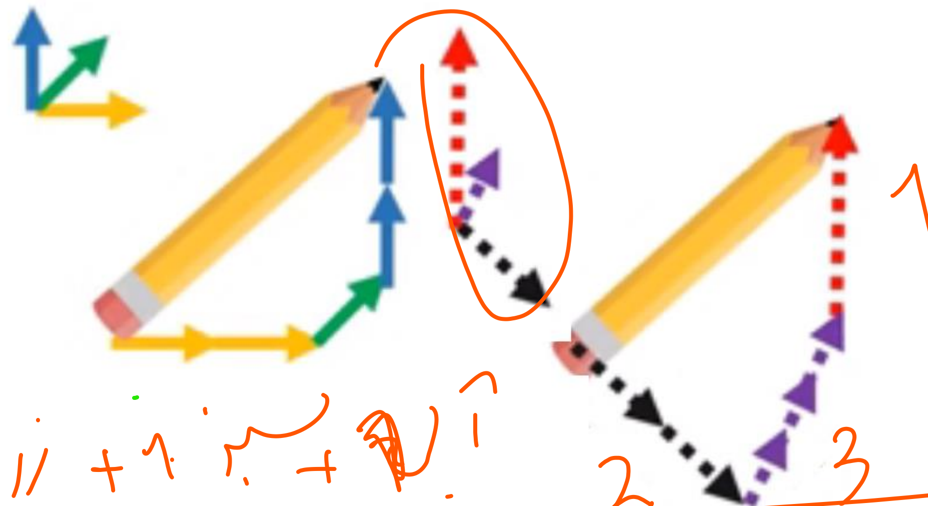
Scalar	Vector	Matrix	Rank-3 Tensor
(rank 0)	(rank 1)	(rank 2)	(rank 3)

Tensor = an object that is invariant under a change of coordinates, and…
has components that change in a special, predictable way under a change of coordinates

Tensor

- Example!





Tensor = a collection of vectors and covectors combined together using the tensor product

Extra: Tensors as partial derivatives and gradients that transform with the Jacobian Matrix

Tensor Product

The $\otimes$ symbol…

Tensor Product (Tensors)	Kronecker Product (Arrays)
$\vec{e}_i \otimes \epsilon^j$	$\begin{bmatrix} v^1 \\ v^2 \end{bmatrix} \otimes \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix}$
Combines 2 tensors into a new 3rd tensor.	Combines 2 arrays into a new 3rd array.

$\vec{e}_i$
vector

ϵ^j
covector

$$\vec{e}_i,\ \epsilon^j \rightarrow \otimes \rightarrow \vec{e}_i \otimes \epsilon^j$$

linear map

Tensor Product

$v = \begin{bmatrix} v^1 \\ \vdots \\ v^i \end{bmatrix}$

$$(\vec{e_i} \otimes \epsilon^j)(\vec{v})$$

covector

$\cancel{\Sigma}\, v^i e_i$

$$= e_i \otimes (\epsilon^j(v)) = e_i \otimes (\epsilon^j(v^k e_k))$$

$$= v^k e_i \otimes (\underbrace{\epsilon^j(e_k)}_{\delta^j_k})$$

$$= v^k e_i \delta^j_k \longrightarrow v^j e_i$$

Tensor vs Kronecker Product

Tensor Product
(Tensors)

$$\vec{v} \otimes \alpha = (v^i \vec{e_i}) \otimes (\alpha_j \epsilon^j)$$

$$= \boxed{v^i \alpha_j} (\vec{e_i} \cancel{\otimes} \epsilon^j)$$

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

2×2

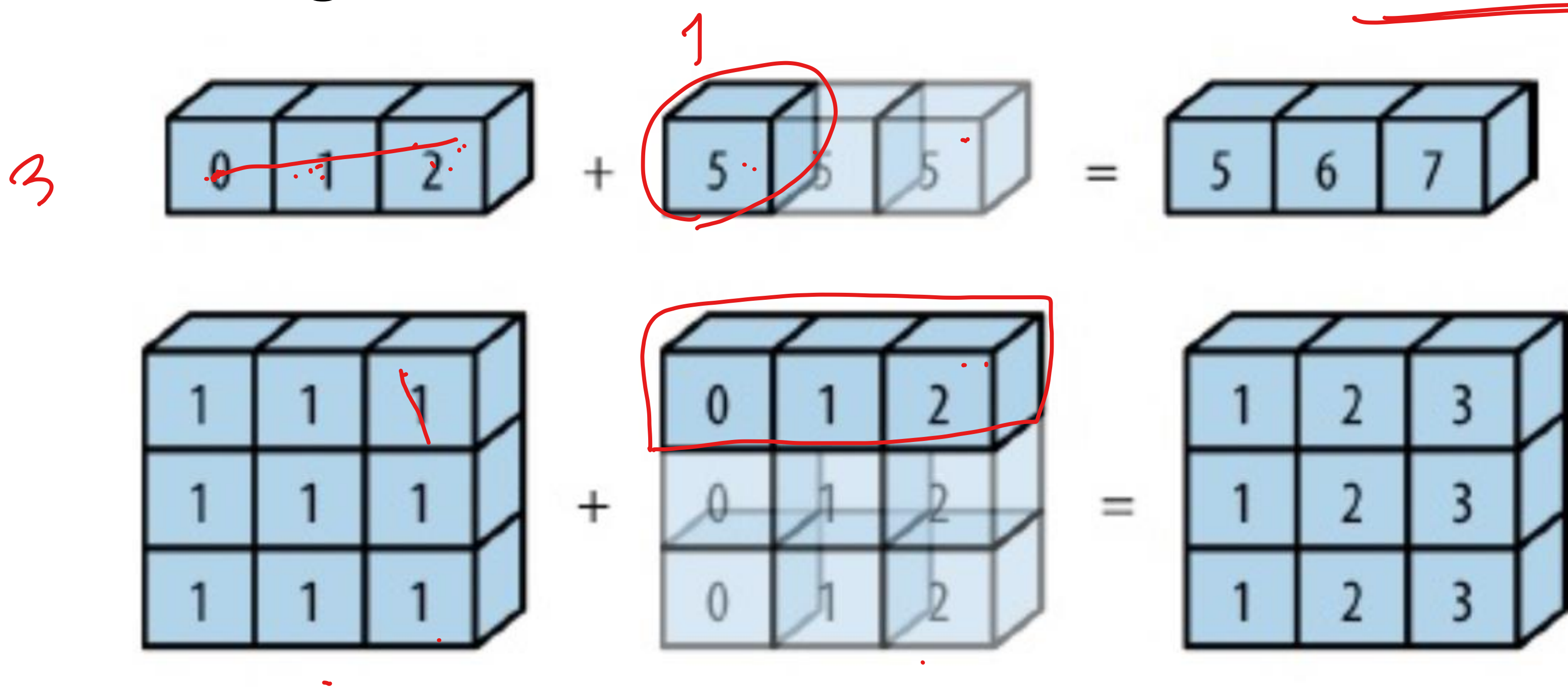
Kronecker Product
(Arrays)

$$\begin{bmatrix} v^1 \\ v^2 \end{bmatrix} \otimes \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix}$$

$$\left[\begin{bmatrix} v^1\alpha_1 \\ v^2\alpha_1 \end{bmatrix} \begin{bmatrix} v^1\alpha_2 \\ v^2\alpha_2 \end{bmatrix} \right]$$

Tensors Addition

- Adding tensors with same size
- Adding scalar to tensor
- Adding tensors with different size: if broacastable



Tensors Addition

- Two tensors are "broadcastable" if the following rules hold:
  - Each tensor has at least one dimension.
  - When iterating over the dimension sizes, starting at the trailing dimension, the dimension sizes must either be equal, one of them is 1, or one of them does not exist.
- Example
  - T1: (5,7,3) T2:(5,7,3)
  - T1: (5,3,4,1) T2:(3,1,1)

Notes

- Matrix Product on tensors

@

$$(m \times n) \cdot (n \times k) = (m \times k)$$

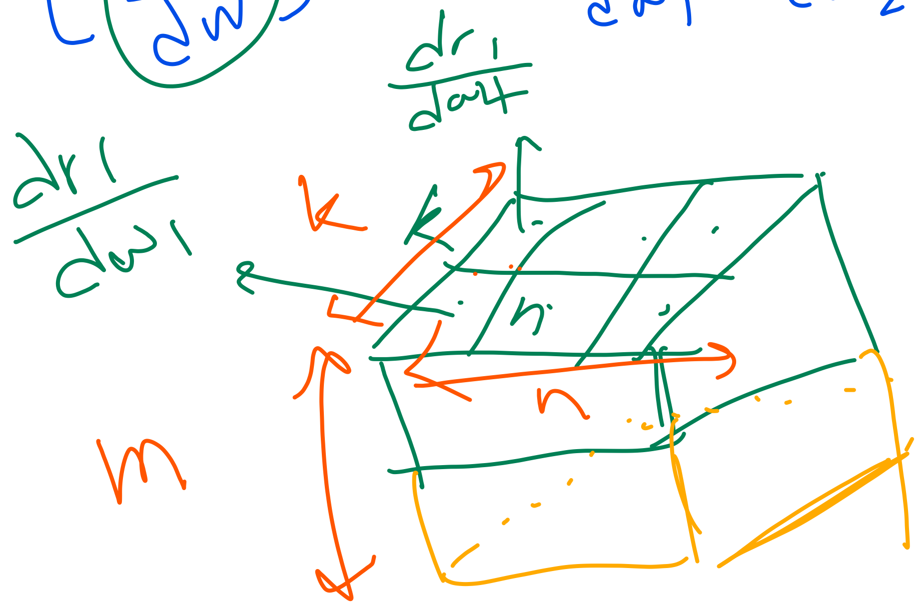
product is defined

Derivative of a vector with respect to a matrix

$$r = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}_{m\times 1} \quad W = \begin{bmatrix} w_1 & w_2 & w_3 \\ w_4 & w_5 & w_6 \end{bmatrix}_{n\times k}$$

$$\frac{dr}{dW} = \begin{bmatrix} \frac{dr_1}{dW} \\ \frac{dr_2}{dW} \end{bmatrix} = \begin{bmatrix} \frac{dr_1}{dw_1} & \frac{dr_1}{dw_2} & \cdots & \frac{dr_1}{dw_6} \\ \frac{dr_2}{dw_1} & \frac{dr_2}{dw_2} & \cdots & \frac{dr_2}{dw_6} \end{bmatrix}$$

$$m \times (nk)$$



$$\underbrace{m \times n \times k}$$

Derivative of a matrix with respect to a matrix

$$A = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \quad m \times n$$

$$\frac{dA}{dW} = \begin{bmatrix} \frac{da_1}{dw} & \frac{da_2}{dw} \\ \frac{da_3}{dw} & \frac{da_m}{dw} \end{bmatrix}$$

$$W = k \times l$$