

# BIOMATH 208 Week 2

Aidan Jan

January 13, 2025

## Multilinear Algebra

### Motivation

two very different types of vectors arise in imaging:

- Tangents to curves/surfaces: Important for describing geometry and anatomy
- Image gradients: Important for optimization and feature identification.

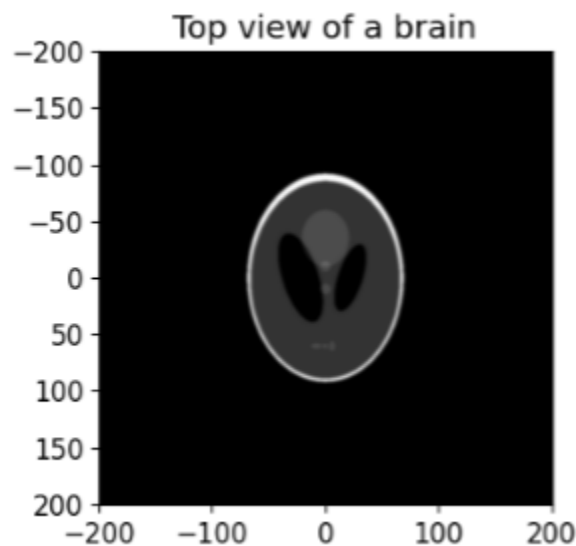
If we chose a new coordinate system to represent our data, these vectors transform differently.

Next, we will apply a linear transform, a rotation matrix  $A$ , to some imaging data and see how these objects change.

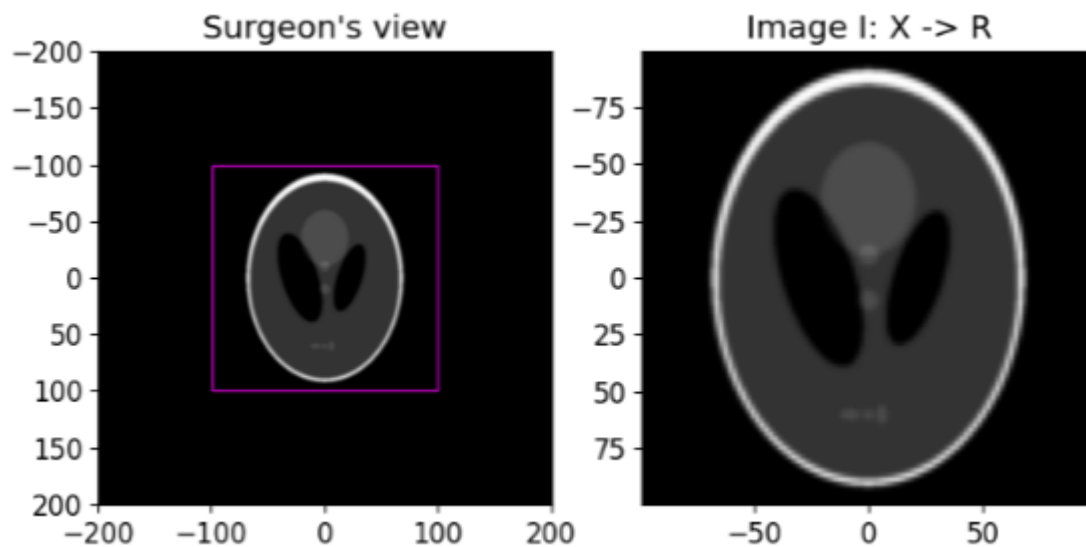
Then we'll apply a different linear transform, a rotation and scale  $B$ , to some imaging data, and see how the objects change.

### Example: Brain Image

Consider this brain image represented by ellipses:

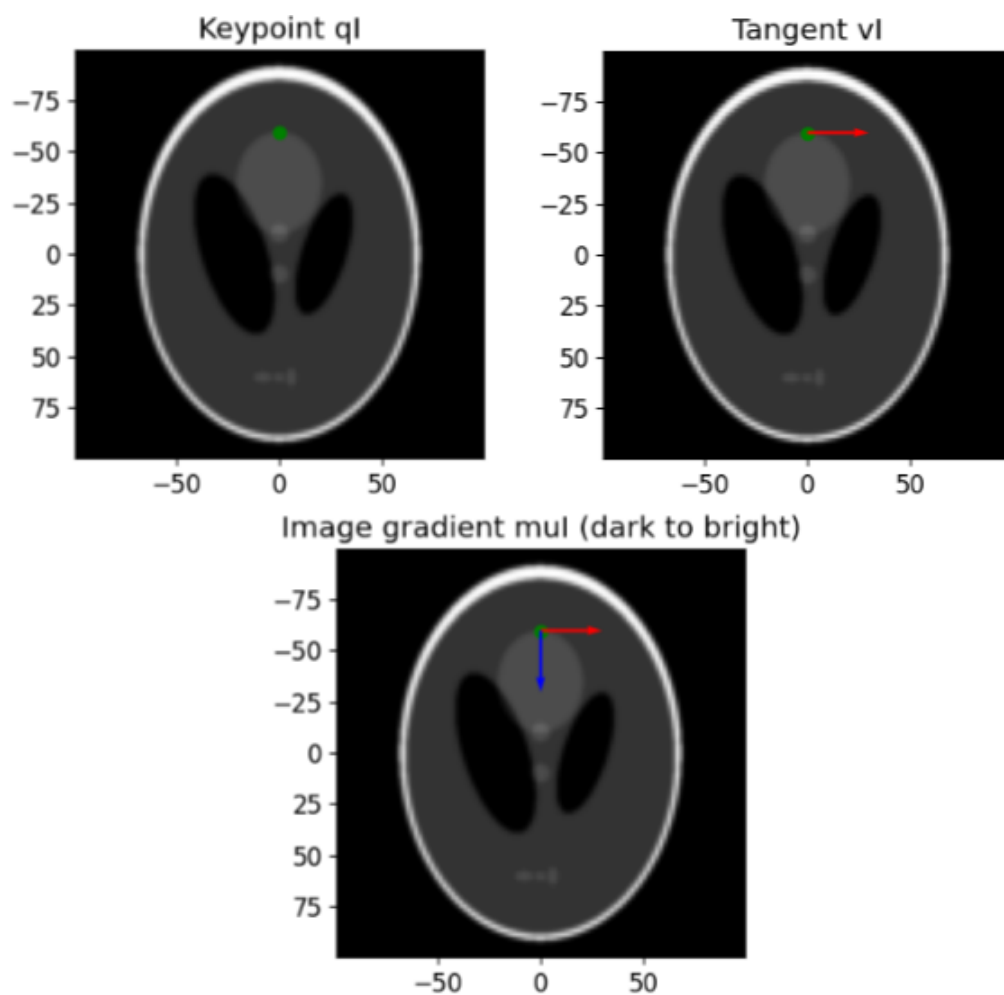


Now, assume this is an image of a patient's brain, and a surgeon is standing directly over it.



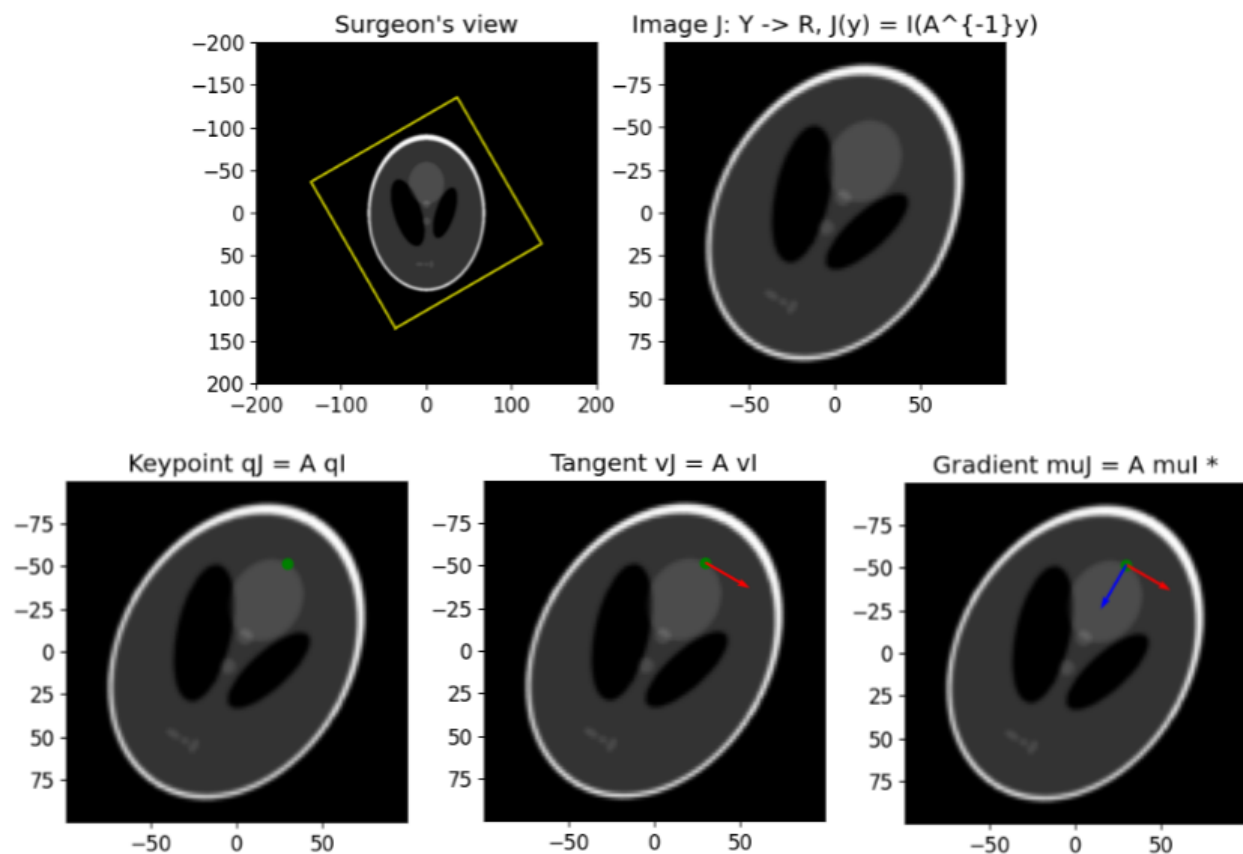
- This is like zooming into the image, thus, Image  $I: X \rightarrow R$

With this transformation, image gradients, tangents, and points are all preserved.



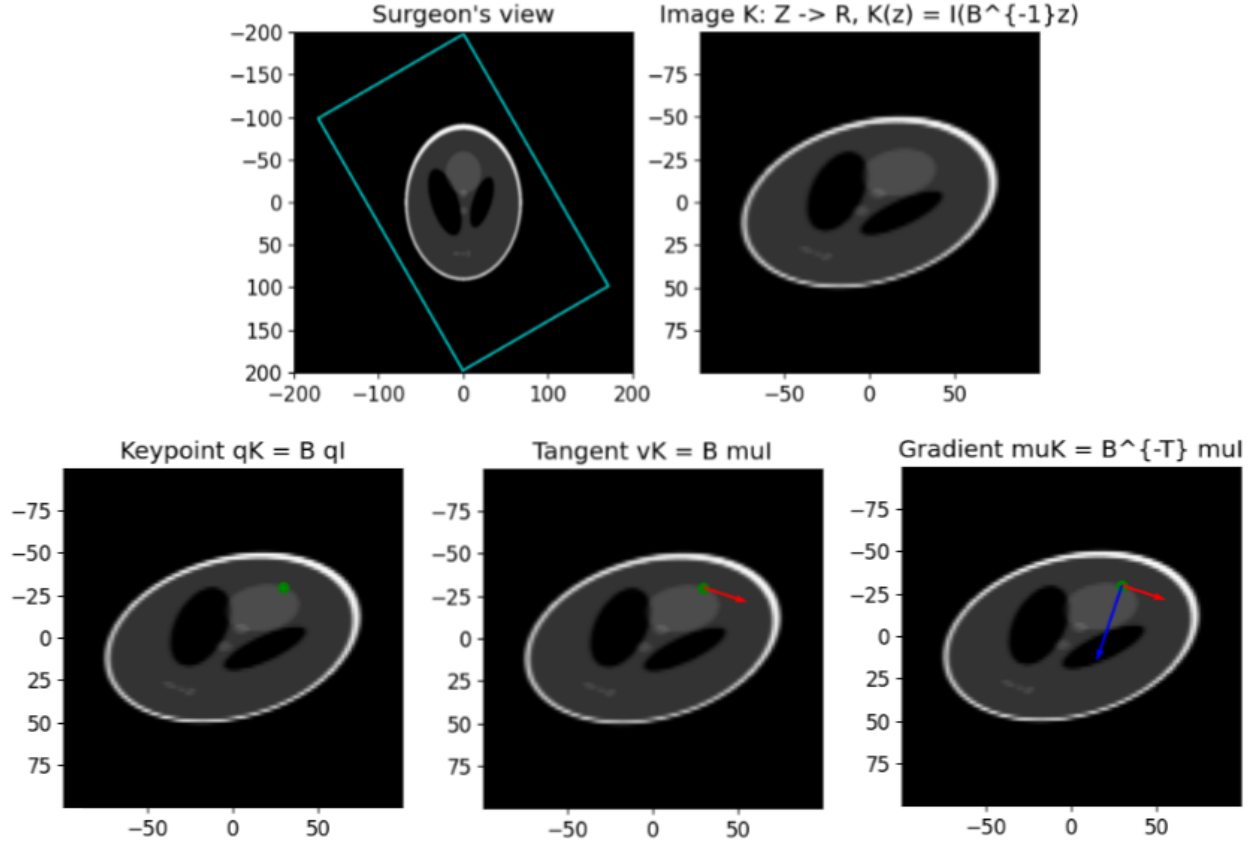
Now, suppose the surgeon turns his head, so he is viewing the head at a different angle:

- A point in the domain of  $I$  is rotated to a point in the domain of  $J$  with the linear map  $A : X \rightarrow Y$ .



In this case, points and vectors are moved based on the rotation matrix multiplied. However, this can be expressed in linear algebra as a change of basis, and thus, points and vectors are preserved relative to the new coordinate system.

Finally, consider if the surgeon moved his head lower, so now he is not looking at the patient's brain from straight top-down, but rather at an angle.



Now, notice that the point may be at the same place relative to the new basis, but the gradient and tangent vectors have shifted.

## Explanation

A tangent  $v$  connects two nearby points,  $q_0, q_1$ . We can define  $v = q_1 - q_0$ . If we transform the points,  $q \rightarrow Aq$ , we transform the vector.

$$Aq_1 - Aq_0 = A(q_1 - q_0) = Av$$

A gradient points from dark to bright in an image. If we transform the image,  $I(x) \rightarrow I(A^{-1}x)$ , we transform the gradient. Let the gradient at  $q$ ,  $g = \nabla l(q)$  be a column vector, then the chain rule says:

$$\nabla[I(A^{-1}x)]_{x=q} = A^{-T} \nabla I(A^{-1}x)_{x=q}$$

so that the transformed gradient is  $A^{-T}g$ .

## Details

What is the  $k$ -th component of  $\nabla I(A^{-1}x)$ ? Let  $B = A^{-1}$  and the components of  $B$  be written as  $B_{ij}$ .

Let  $(x, y)$  map to  $(B_{00}x + B_{01}y, B_{10}x + B_{11}y)$ . To find the gradient, we take the derivative.

$$\frac{\partial}{\partial x} I(B_{00}x + B_{01}y, B_{10}x + B_{11}y)$$

By chain rule, ...

$$= \partial_0 I(B_{00}x + B_{01}y, B_{10}x + B_{11}y) B_{00} + \partial_1 I(B_{00}x + B_{01}y, B_{10}x + B_{11}y) B_{10}$$

$\partial_0$  means "partial derivative with respect to the 0th argument"

Now, we take the derivative in terms of  $y$ .

$$\begin{aligned} \frac{\partial}{\partial y} I(B_{00}x + B_{01}y, B_{10}x + B_{11}y) \\ = \partial_0 I(B_{00}x + B_{01}y, B_{10}x + B_{11}y) B_{01} + \partial_1 I(B_{00}x + B_{01}y, B_{10}x + B_{11}y) B_{11} \end{aligned}$$

What we have essentially done here is:

$$\begin{aligned} \begin{pmatrix} B_{00} & B_{10} \\ B_{01} & B_{11} \end{pmatrix} \cdot \begin{pmatrix} \partial_0 I(B_{00}x + B_{01}y, B_{10}x + B_{11}y) \\ \partial_1 I(B_{00}x + B_{01}y, B_{10}x + B_{11}y) \end{pmatrix} \\ = B^T \cdot \nabla I(B_{00}x + B_{01}y, B_{10}x + B_{11}y) \\ = A^{-T} \nabla I(B_{00}x + B_{01}y, B_{10}x + B_{11}y) \end{aligned}$$

$A^{-T}$  is the inverse and transpose of  $A$ . This is equal to  $B^T$ , since by our definition,  $B = A^{-1}$ .

## Definition of Vector Space

A set  $V$  with two operations  $+$  (plus) and  $\cdot$  (times, i.e., scalar multiplication) is a vector space if the operations satisfy CANI-ADDU. Here, let  $u, v, w \in V$  and  $a, b \in \mathbb{R}$ .

1. **Commutativity:**  $u + v = v + u$
2. **Associativity:**  $(u + v) + w = u + (v + w) = u + v + w$ 
  - $+$  needs to be a function of two variables
3. **Neutral element:**  $\exists O \in V$  such that  $O + u = u + O = u$ 
  - This is often the zero vector, or the origin.
4. **Inverse element:**  $\exists(-v)$  such that  $v + (-v) = (-v) + v = 0$
5. **Associativity:**  $a \cdot (b \cdot u) = (a \cdot b) \cdot u$ , where  $(b \cdot u) \rightarrow \mathbb{R} \times V \rightarrow V$ , and  $(a \cdot b) \rightarrow \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$
6. **Distributivity over  $+$  on  $V$ :**  $a \cdot (u + v) = a \cdot u + a \cdot v$
7. **Distributivity over  $\cdot$  on  $\mathbb{R}$ :**  $(a + b) \cdot u = a \cdot u + b \cdot u$
8. **Unitary element:**  $\exists 1$  such that  $1 \cdot u = u$ .

## Example Vector Spaces

### Example: Real Numbers

Real numbers form a vector space with the usual definition of  $+$  and  $\cdot$ .

- a single number, not pairs or triples
- not "direction and magnitude" but still obey vector space axioms

### Example: The Cartesian Plane

Set of ordered pairs,  $\mathbb{R} \times \mathbb{R}$ .  $(a, b) \rightarrow$  a vector in the cartesian plane.

We can define  $+$  and  $\cdot$  as the following:

- $+$ :  $(a, b) + (c, d) = (a + c, b + d)$ , where all  $a, b, c, d, a + c, b + d \in \mathbb{R}$ .
- $\cdot$ :  $c \cdot (a, b) = (c \cdot a, c \cdot b)$ , where  $a, b, c, c \cdot a, c \cdot b \in \mathbb{R}$ .

Using these rules, we can build a more complex vector space, using definitions from a simpler one.

**Example: Arrows on paper with tail at a fixed origin**

A vector is an arrow you can draw on paper. We can define  $+$  and  $\cdot$  as the following:

- $+$ : "translate the second arrow so its tail lines up with the tip of the first arrow, draw a new arrow from the origin to the tip of the second." Basically, add the vectors.
- $\cdot$ : "shrink or grow the vector by this factor, and if the factor is negative, reflect it over the origin"

**Example: Functions from  $\mathbb{R} \rightarrow \mathbb{R}$** 

Let  $f, g$  be functions  $\rightarrow$  choose a set of functions with properties we like

- Define  $f + g$  by evaluating any point  $x \in \mathbb{R}$ , define  $(f + g)(x) = f(x) + g(x)$ .  $f$  and  $g$  are functions  $\mathbb{R} \rightarrow \mathbb{R}$ .
- Define  $a \cdot f$  by evaluating at any point  $x \in \mathbb{R}$ , define  $(a \cdot f)(x) = a \cdot f(x)$ , where  $(a \cdot f \in V)$  and  $a \cdot f(x) \in \mathbb{R}$ .

**Example: Probabilities**

Let  $p$  represent the probability that an event happens, and  $q$  represent the probability that another event happens.

- $+$ :  $p + q$  might be bigger than 1, which is not a probability!
- $\cdot$ :  $(-1) \cdot p$  is less than 0, which is also not a probability!

However, if we use logits,  $\log \frac{p}{1-p}$ , we can define  $+$  and  $\cdot$  without any problems. (This will be covered at the end of the course.)

**Example: Arrows anywhere on paper**

The idea with this one is that we do not necessarily have a fixed origin. Because of this, we do not have a good definition of  $+$  and  $\cdot$ . Or at least, that is when we imagine the vectors are on a flat sheet of paper.

Consider if the vectors were drawn on a sphere.

- If two vectors are tangent to the same point, we can add them tip to tail.
- If they are not, then there is no way to add the vectors. We would consider them to be vectors in different vector spaces.

**Linear Algebra Definitions**

- Linear Independence: A set of vectors  $v_i$  are linearly independent if

$$\sum a_i v_i = 0 \iff a_i = 0 \forall i$$

Otherwise they are linearly dependent.

- Basis: A linear independent set of vectors  $B$  is called a basis if any vector in  $V$  can be written as a unique linear combination of its elements.
- Dimension: The number of vectors in this set is called the dimension of the vector space.
- Coordinates: If the dimension is finite, we write basis elements as  $e_i$ , then

$$v = \sum_i v^i e_i$$

for  $v^i \in \mathbb{R}$ . The  $v^i$  are the coordinates of the vector with respect to this basis.

- Einstein summation notation: In expressions like the above we will usually remove the summation sign and write

$$v = v^i e_i$$

Any pair of indices with one up and one down is summed.

## Color Convention

We will use the convention

- vectors in red
- scalars in black

The previous example will be written

$$v = v^i e_i$$

## Example: Polynomials

### Linear Maps

A linear map  $L : U \rightarrow V$ , is a map between two vector spaces  $(U, V)$  that is compatible with addition and scalar multiplication. For  $u, v \in U$ , and  $a \in \mathbb{R}$ , this means:

1.  $+$ :  $L(u + v) = L(u) + L(v)$ 
  - "Adding the inputs is the same as adding the outputs"
2.  $\cdot$ :  $L(a \cdot u) = a \cdot L(u)$

**Linear Maps are a vector space.**

**Definition: The vector space of linear maps:**

The set of linear maps from  $U$  to  $V$  is a vector space,  $\text{Hom}(U, V)$  (homomorphisms). We define  $+$  and  $\cdot$  for each input, borrowing the definition from  $V$ . Let  $\text{Hom}(U, V) = H$ ,  $A, B \in H$ ,  $u \in U$ , and  $a \in \mathbb{R}$ .

- $+$ :  $A +_H B$ , where  $+_H$  is a homomorphism,  $(A +_H B)(u) = A(u) +_V B(u)$
- $\cdot$ :  $a \cdot_H A$ ,  $(a \cdot_H A)(u) = a \cdot_V A(u)$
- e.g., we can add matrices elementwise, and we can scale them elementwise.

**Definition: Endomorphism:**

The set of linear maps  $\text{Hom}(V, V)$  are called endomorphisms.

## Covectors

A covector is the vector space  $\text{Hom}[V, \mathbb{R}]$ , denoted  $V^*$  are called covectors.  $V^*$  is called the dual space (or just dual) of  $V$ .  $V$  is called primal.

### Example: Row and Column Vectors

Often we write the components of vectors in  $V$  as a column. Then the components of  $V^*$  are written as a row, and act on the left using matrix multiplication.

### Example: Derivatives

Partial derivatives of a function are covectors, and are often written as a row.

## Color Convention Pt. 2

We will expand our convention:

- vectors in red
- covectors in blue
- scalars in black.

Vectors will generally be denoted with Roman letters near the end of the alphabet (e.g.,  $u, v, w, x, y, z$ ).

Covectors will generally be denoted with Greek letters (e.g.,  $\mu, \nu, \omega, \chi$ )

Scalars will generally be denoted with Roman letters near the beginning of the alphabet (e.g.,  $a, b, c$ ).

## A Basic for Covectors

### Definition: Dual Basis:

Once a basis for  $V$  is fixed ( $e_i$ ), we can uniquely define a basis of  $V^*$ ,  $\epsilon^i$  by the relationship

$$\epsilon^i(e_j) = \delta_j^i = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

## Recovering Vector Coordinates

If a vector is written in coordinates as  $v = v^i e_i$ , then the  $i$ -th coordinate is found by

$$v^i = \epsilon^i(v)$$

**Proof:**

$$\begin{aligned} \epsilon^i(v) &= \epsilon^i(v^j e_j) \\ &= v^j \cdot \epsilon^i(e_j) \\ &= v^j \delta_j^i && \text{Definition of dual basis} \\ &= v^i && \text{All terms are 0 except for the } i\text{-th term} \end{aligned}$$

## Recovering Covector Coordinates

If a covector is written in coordinates as  $\mu = \mu^i \epsilon^i$ , then the  $i$ -th coordinate is found by:

$$\mu^i = \mu(e_i)$$

**Proof:**

$$\begin{aligned} \mu(e_i) &= (\mu_j \epsilon^j)(e_i) \\ &= \mu_j \epsilon^j(e_i) \\ &= \mu_j \delta_i^j && \text{Definition of dual} \\ &= \mu_i \end{aligned}$$

## Dual of the dual

The vector space of linear maps from  $V^* \rightarrow \mathbb{R}$  is denoted  $V^{**}$ , and is isomorphic to  $V$ .

If  $\mu \in V^*$ , and  $w \in V^{**} \simeq V$ , then we define

$$w(\mu) \doteq \mu(w)$$

We may use this notation when we consider a sequence of maps acting from right to left. Often we will leave out parentheses.



### Example: Dual Basis for Ordered Pairs

Let  $V = \mathbb{R}^2$ , and a basis be given by  $e_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  and  $e_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ . Considering  $V^*$  as row vectors, find the dual basis.

Essentially, we are solving for  $\epsilon^0$  and  $\epsilon^1$ .

$$\begin{aligned}\epsilon^0(e_0) &= \delta_0^0 \\ &= 1 = \epsilon^{00} \cdot 1 + \epsilon^{01} \cdot 0 \\ &= \epsilon^{00}\end{aligned}$$

The multiplication by 1 and 0 are the components of  $e_0$ .

$$\begin{aligned}\epsilon^0(e_1) &= \delta_1^0 \\ &= 0 = \epsilon^{00} \cdot 1 + \epsilon^{01} \cdot 1 \\ &= \epsilon^{00} + \epsilon^{01} = 0\end{aligned}$$

In this case, we multiply by 1 for both  $\epsilon$  terms since the components of  $e_1$  are both 1.

$$\begin{aligned}\epsilon^1(e_0) &= \delta_0^1 \\ &= 0 = \epsilon^{10} \cdot 1 + \epsilon^{11} \cdot 0 \\ &= \epsilon^{10}\end{aligned}$$

$$\begin{aligned}\epsilon^1(e_1) &= \delta_1^1 \\ &= 1 = \epsilon^{10} \cdot 1 + \epsilon^{11} \cdot 1 \\ &= \epsilon^{10} + \epsilon^{11} = 1\end{aligned}$$

At this point, we have four linear equations and four variables,  $\epsilon^{00}, \epsilon^{01}, \epsilon^{11}, \epsilon^{10}$ . Solving, we get:

$$\begin{aligned}\epsilon^0 &= (1, -1) \\ \epsilon^1 &= (0, 1)\end{aligned}$$