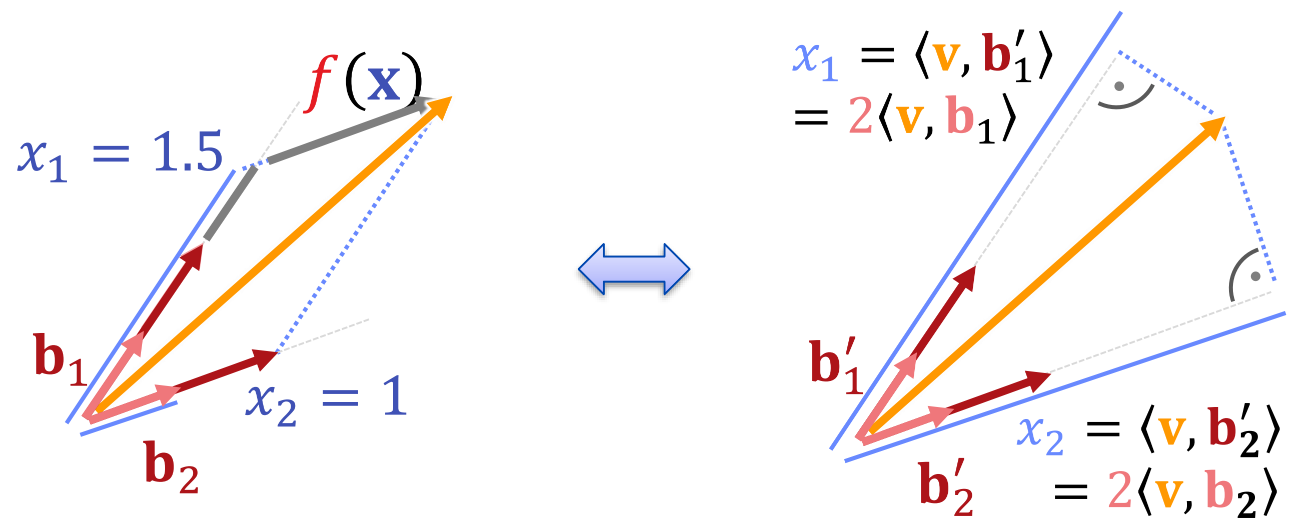
Modelling 1

SUMMER TERM 2020

JOHANNES GUTENBERG UNIVERSITÄT MAINZ



ADDENDUM

Co- and Contravariance

Michael Wand · Institut für Informatik · michael.wand@uni-mainz.de

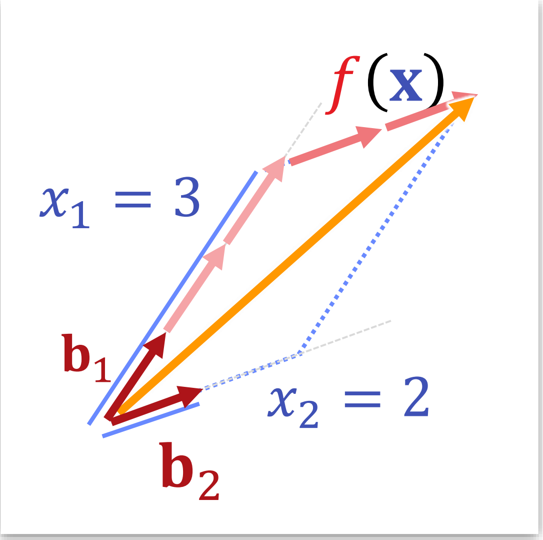
Covariance & Contravariance

Representing Vectors

Two operations in linear algebra

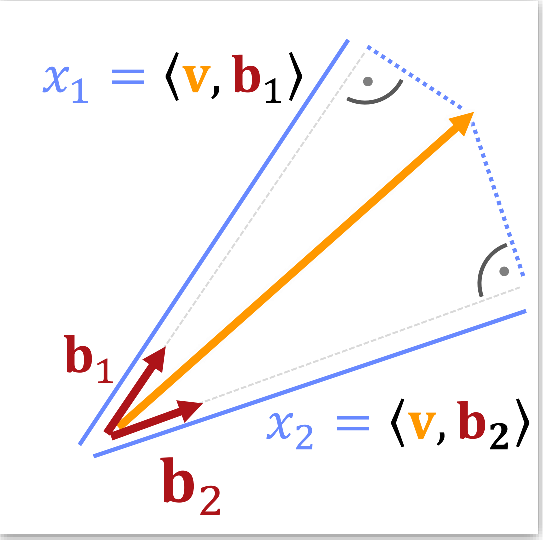
- **Contravariant:**
  Linear combination of vectors

$$\mathbf{v} = \sum_{i=1}^{n} x_i \mathbf{b}_i \quad \rightarrow \quad \mathbf{v} \equiv \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$



- **Covariant:**
  Projection on vectors (w/scalar product)

$$\mathbf{v} \equiv \begin{pmatrix} \langle \mathbf{v}, \mathbf{b}_1 \rangle \\ \vdots \\ \langle \mathbf{v}, \mathbf{b}_n \rangle \end{pmatrix}$$



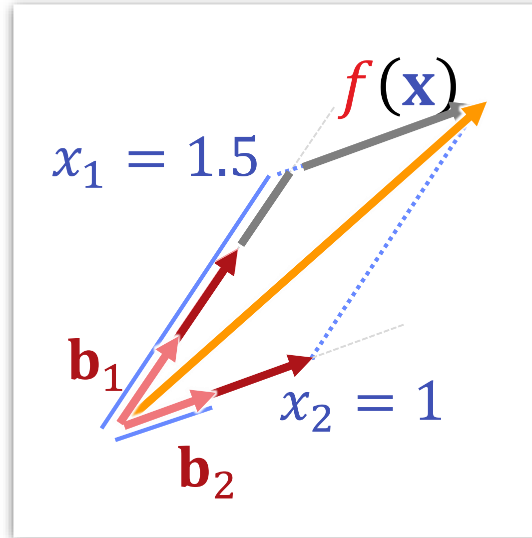
Where is the difference?

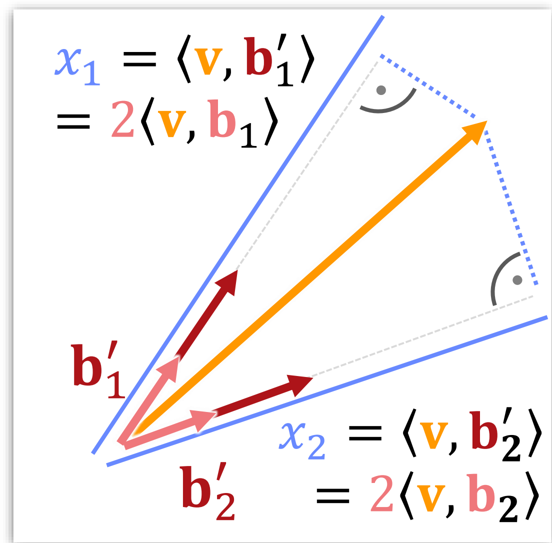
Change of basis

- **Contravariant:** $f(\mathbf{x}) = \sum_{i=1}^{n} x_i \mathbf{b}_i$
- Keep same output vector:
  $\mathbf{b}_i \rightarrow \mathbf{T}\mathbf{b}_i$ *requires* $\mathbf{x} \rightarrow \mathbf{T}^{-1}\mathbf{x}$
- **Covariant:**

$$f(\mathbf{x}) = \begin{pmatrix} \langle \mathbf{x}, \mathbf{b}_1 \rangle \\ \vdots \\ \langle \mathbf{x}, \mathbf{b}_n \rangle \end{pmatrix}$$

- Keep same output vector:
  $\mathbf{b}_i \rightarrow \mathbf{T}\mathbf{b}_i$ *requires* $\mathbf{x} \rightarrow \mathbf{T}\mathbf{x}$





Awesome Video

Tensors, Co-/Contra-Variance

- „Tensors Explained Intuitively: Covariant, Contravariant, Rank“
  Physics Videos by Eugene Khutoryansky

  https://www.youtube.com/watch?v=CliW7kSxxWU

Covariance & Contravariance

Linear map

$$\mathbf{f}: V_1 \to V_2$$

Matrix representation (standard basis)

$$\mathbf{M} \in \mathbb{R}^{d_1 \times d_2}$$

Change of basis

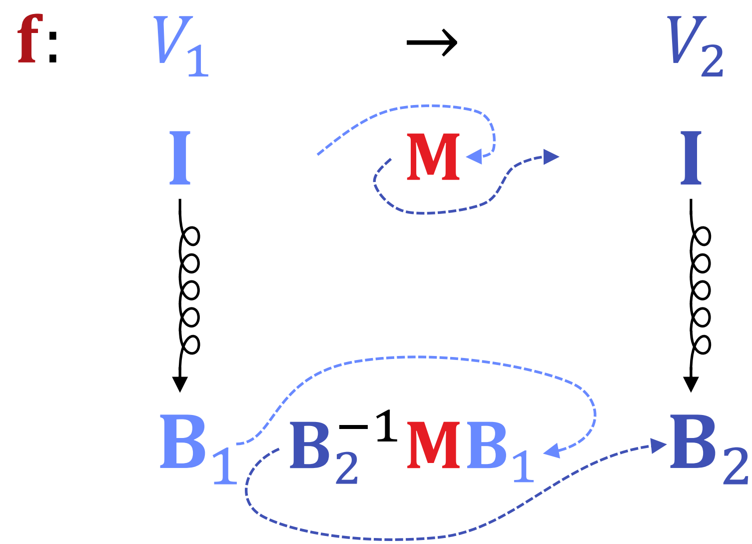
$$\mathbf{B}_1 = \begin{pmatrix} | & & | \\ \mathbf{b}_1^{(1)} & \cdots & \mathbf{b}_{d_1}^{(1)} \\ | & & | \end{pmatrix}, \qquad \mathbf{B}_2 = \begin{pmatrix} | & & | \\ \mathbf{b}_1^{(2)} & \cdots & \mathbf{b}_{d_2}^{(2)} \\ | & & | \end{pmatrix}$$

New matrix representation (bases $\mathbf{B}_1$, $\mathbf{B}_2$)

$$\mathbf{B}_2^{-1}\mathbf{M}\mathbf{B}_1 \in \mathbb{R}^{d_1 \times d_2}$$

Covariance & Contravariance

Situation

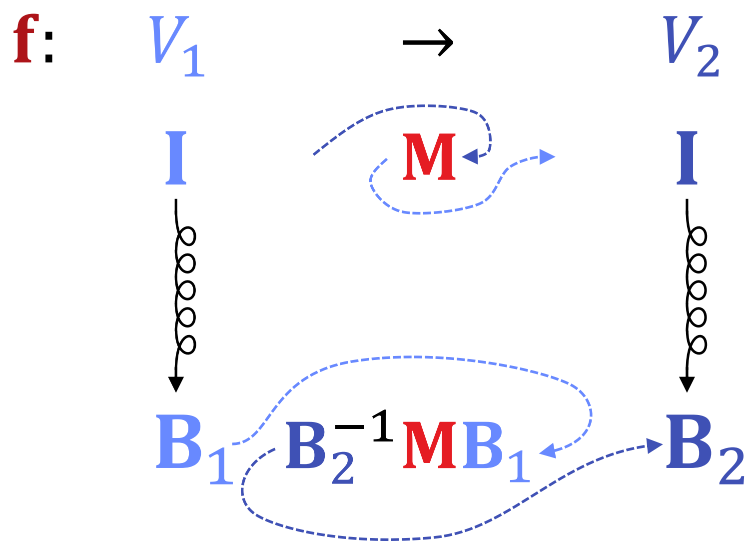


Transformation law

- Input vectors $\mathbf{x}$ $(\mathbf{M}\mathbf{x})$: $\mathbf{x}_{[\mathbf{B}_1]} = \mathbf{B}_1 \mathbf{x}_{[\mathbf{I}]}$ *(covariant)*
- Output vectors $\mathbf{y} = \mathbf{M}\mathbf{x}$: $\mathbf{y}_{[\mathbf{B}_2]} = \mathbf{B}_2^{-1} \mathbf{y}_{[\mathbf{I}]}$ *(contravariant)*

Covariance & Contravariance

Situation



Transformation law

- Input vectors $\mathbf{x}$ $(\mathbf{M}\mathbf{x})$: $\mathbf{x}_{[\mathbf{B}_1]} = \mathbf{B}_1 \mathbf{x}_{[\mathbf{I}]}$ *(covariant)*
- Output vectors $\mathbf{y} = \mathbf{M}\mathbf{x}$: $\mathbf{y}_{[\mathbf{B}_2]} = \mathbf{B}_2^{-1} \mathbf{y}_{[\mathbf{I}]}$ *(contravariant)*

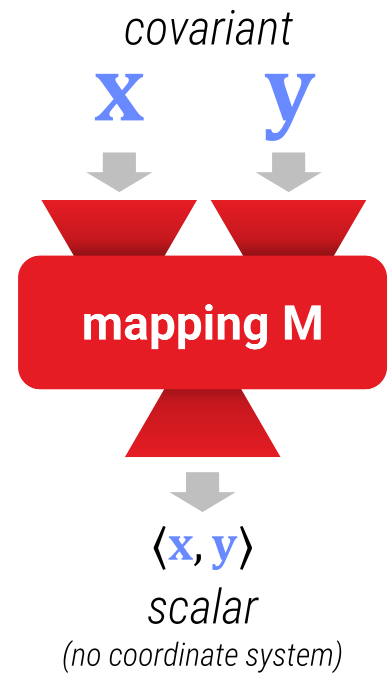
Covariance & Contravariance

$$f(\mathbf{x}) \leftarrow \mathbf{B}_2^{-1}\mathbf{M}\mathbf{B}_1 \leftarrow \mathbf{x}$$

$$\mathbf{B}_2^{-1}\underbrace{(\mathbf{M}\mathbf{B}_1)}_{\text{transforms row-vectors}}$$

$$\underbrace{(\mathbf{B}_2^{-1}\mathbf{M})}_{\text{transforms column-vectors}}\mathbf{B}_1$$

Scalar Product



General scalar product

$$\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$$

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^{\mathrm{T}} \mathbf{Q} \, \mathbf{y},$$

$$\left(\mathbf{Q} = \mathbf{Q}^{\mathrm{T}}, \mathbf{Q} > 0 \right)$$

Change of basis

$$\mathbf{I} \rightsquigarrow \mathbf{B}$$

$$\mathbf{x} \to \mathbf{B}\mathbf{x}, \quad \mathbf{y} \to \mathbf{B}\mathbf{y}$$

$$\langle \mathbf{x}, \mathbf{y} \rangle_{[\mathbf{I}]} = \mathbf{x}^{\mathrm{T}} \mathbf{Q} \, \mathbf{y}$$

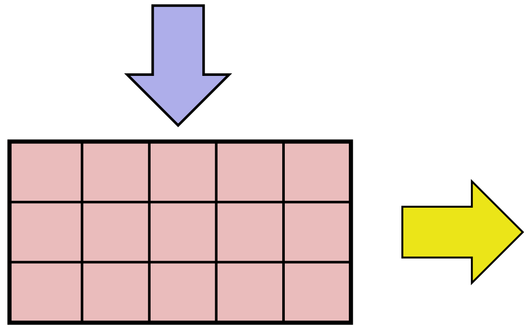
$$\langle \mathbf{x}, \mathbf{y} \rangle_{[\mathbf{B}]} = \mathbf{x}^{\mathrm{T}} \cdot [\mathbf{B}^{\mathrm{T}} \cdot \mathbf{Q} \cdot \mathbf{B}] \cdot \mathbf{y}$$

Three shades of dual

PCA, SVD, MDS

Inputs and Outputs

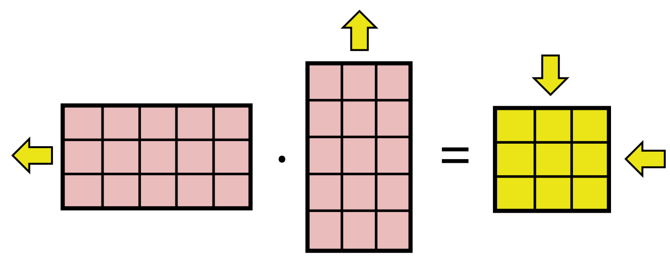
Input ("covariant") side
of the matrix



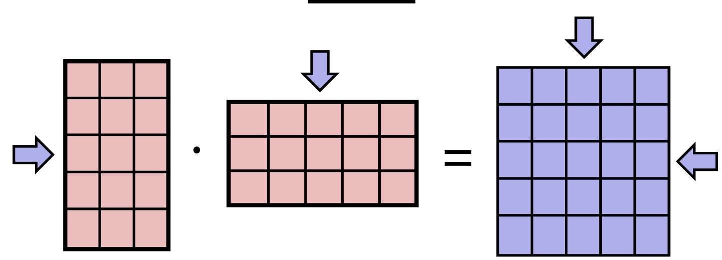
Output ("contravariant")
side of the matrix

Squaring a Matrix

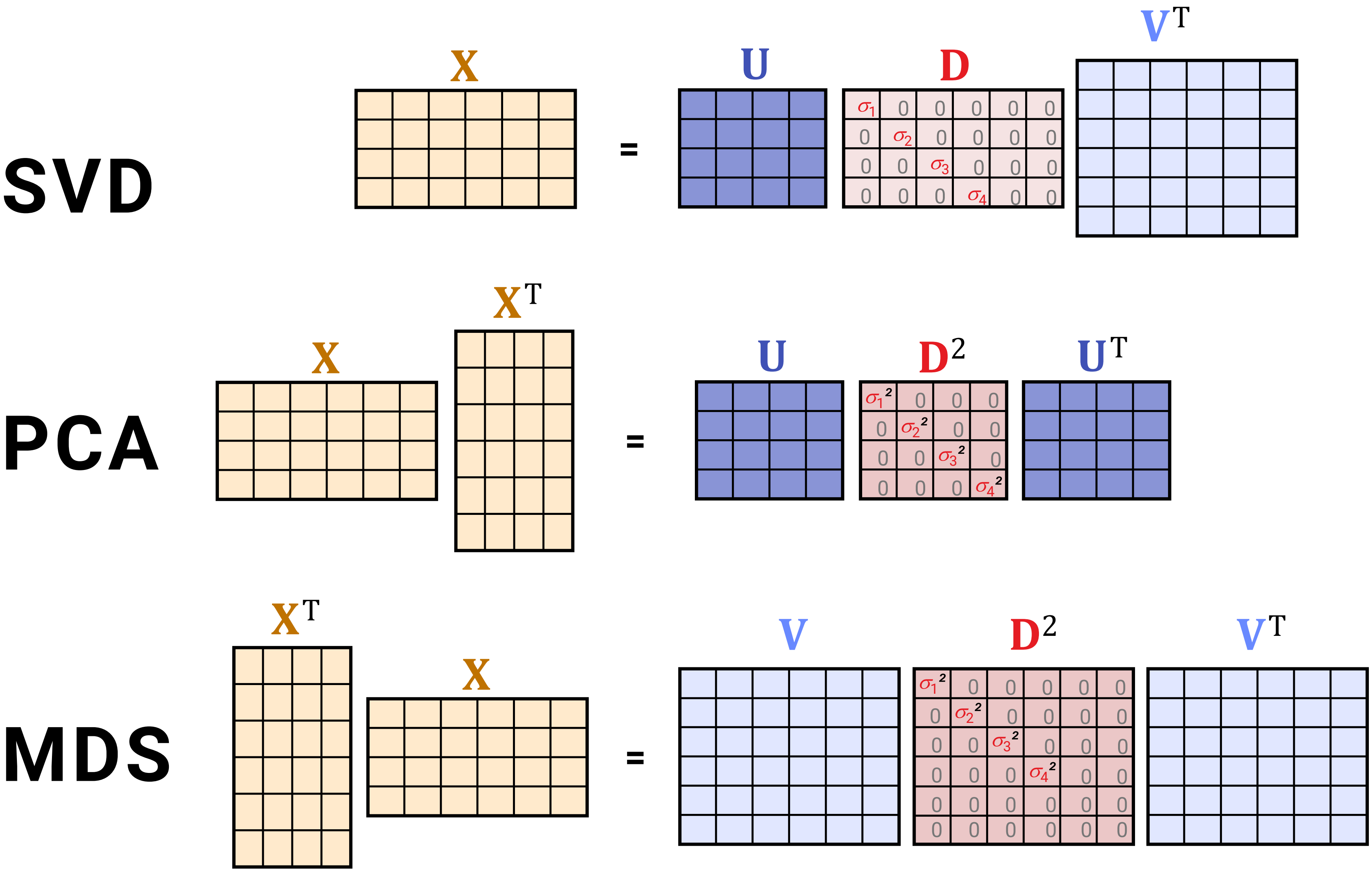
- Possibility 1: $\mathbf{A} \cdot \mathbf{A}^T$



- Possibility 2: $\mathbf{A}^T \cdot \mathbf{A}$



A Story about Dual Spaces



Tensors:
Multi-Linear Maps

Tensors

General notion: Tensor

- Tensor: multi-linear form*) with $r \in \mathbb{N}$ input vectors

  $$\mathbf{T}: V_1, \dots, V_n, V_{n+1}, \dots, V_r \to F$$ (usually: field $F = \mathbb{R}$)

  - "Rank r" tensor
  - Linear in each input (when keeping the rest constant)
- Each input can be covariant or contravariant
  - (n, m) tensor
  - $r = n + m$
  - n – contravariant inputs
  - m – covariant inputs

*) i.e.: multi-linear mapping / function

Tensors

Representation

- Represented as r-dimensional array

$$t^{i_1,i_2,\ldots,i_n}_{j_1,j_2,\ldots,j_m}$$

  - n – contravariant inputs (“indices”)
  - m – covariant inputs (“indices”)

- Mapping rule

$$\mathbf{T}\left(\mathbf{v}^{(1)},\ldots,\mathbf{v}^{(n)},\mathbf{w}^{(1)},\ldots,\mathbf{w}^{(m)}\right) :=$$

$$\sum_{i_1=0,\ldots,n_{i_1}} \cdots \sum_{i_n=0,\ldots,n_{i_n}} \sum_{j_1=0,\ldots,n_{j_1}} \cdots \sum_{j_m=0,\ldots,n_{j_m}} v^{(1)}_{i_1} \cdots v^{(n)}_{i_n} w^{(1)}_{j_1} \cdots w^{(m)}_{j_m} t^{i_1,i_2,\ldots,i_n}_{j_1,j_2,\ldots,j_m}$$

(Note: writing the application of $\mathbf{T}$ as multi-linear mapping here)

Tensors

Remarks

- No difference between co-/contravariant dimensions in terms of numerical representation
- Generalization of matrix

Example

$$\mathbf{T}\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}\right) = 42x_1y_1z_1 + 23x_1y_1z_2 + \cdots + 16x_2y_2z_3$$

- Purely linear polynomial in each input parameter when all others remain constant.
- 3D array - $2 \times 2 \times 3$ combinations of coefficients

Einstein Notation

Example: Quadratic polynomial $\mathbb{R}^3 \to \mathbb{R}^3$

$$p^j(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} + \mathbf{b}\mathbf{x} + \mathbf{c}$$

$$p^j = \left[\sum_{k=1}^{3}\sum_{l=1}^{3} x_k x_l \, a_{kl}^j\right] + \left[\sum_{k=1}^{3} x_k b_k^j\right] + \mathbf{c}^j$$

Tensor notation

- Input: x_i, $i = 1..3$
- Output: p^j, $j = 1..3$
- Quadratic form (Matrix) $\mathbf{A}$: a_{kl}
- Linear form (Co-Vector) $\mathbf{b}$: b_k
- Constant c

Einstein Notation

Example: Quadratic polynomial $\mathbb{R}^3 \to \mathbb{R}^3$

$$p^j(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} + \mathbf{b}\mathbf{x} + \mathbf{c}$$

Einstein notation (implicit sums over common indices)

$$p^j = x_k x_l a_{kl}^j + x_k b_k^j + \mathbf{c}^j$$

Tensor notation

- Input: x_i, $i = 1..3$
- Output: p^j, $j = 1..3$

- Quadratic form (Matrix) $\mathbf{A}$: a_{kl}
- Linear form (Co-Vector) $\mathbf{b}$: b_k
- Constant c

Further Examples

Examples

- (n, m)-tensor
  - n contravariant “indices”
  - m covariant “indices”
- Matrix: $(1,1)$-tensor
- Scalar product: $(0,2)$-tensor
- Vector: $(1,0)$-tensor
- Co-vector: $(0,1)$-tensor
- Geometric vectors: $(1,0)$-tensors

Covariant Derivatives?

Examples

- Geometric vectors: (1,0) tensors
- Derivatives*) / gradients / normal vectors: (0,1) tensors

  *) **to be precise:**

  - Spatial derivatives co-vary for changes of the basis of the space
    - $f: \mathbb{R}^n \rightarrow \mathbb{R},\ f(\mathbf{x}) = \mathbf{y},\ \Rightarrow \nabla f$ is covariant (0,1).
    - Examples: Gradient vector
  - Derivatives of vector functions by unrelated dimensions remain contravariant
    - $f: \mathbb{R} \rightarrow \mathbb{R}^n,\ f(t) = \mathbf{y},\ \Rightarrow \frac{\mathrm{d}}{\mathrm{d}t} f$ remains contravariant (1,0).
    - Examples: velocity, acceleration
  - Mixed case: $f: \mathbb{R}^n \rightarrow \mathbb{R}^m, \nabla f = J_f$ is a $(1,1)$-tensor (1,1)

Example: Plane Equation

Plane equation(s)

- Parametric:

$$\mathbf{x} = \lambda_1 \mathbf{r}_1 + \lambda_2 \mathbf{r}_2 + \mathbf{o}$$

- Implicit:

$$\langle \mathbf{n}, \mathbf{x} \rangle - d = 0$$

Transformation $\mathbf{x} \to \mathbf{T}\mathbf{x}$

- Parametric:

$$\mathbf{T}\mathbf{x} = \mathbf{T}(\lambda_1 \mathbf{r}_1 + \lambda_2 \mathbf{r}_2 + \mathbf{o}) = \lambda_1 \mathbf{T}\mathbf{r}_1 + \lambda_2 \mathbf{T}\mathbf{r}_2 + \mathbf{T}\mathbf{o}$$

- Implicit:

$$\langle \mathbf{n}, \mathbf{T}\mathbf{x} \rangle - d = \left(\mathbf{n}^{\mathrm{T}} \mathbf{T}\right) \mathbf{x} - d0$$

More Structure?

Connecting

- Integrals
- Derivatives
- In higher dimensions
- And their transformation rules

"Exterior Calculus"

- Unified framework
- Beyond this lecture (take a real math course :-))

Vectors & Covectors in Function Spaces

Remark: Function Spaces

Discrete vector spaces

- Picking entries by index is a linear operation
- Can be represented by projection to vector (multiplication with "co-vector")

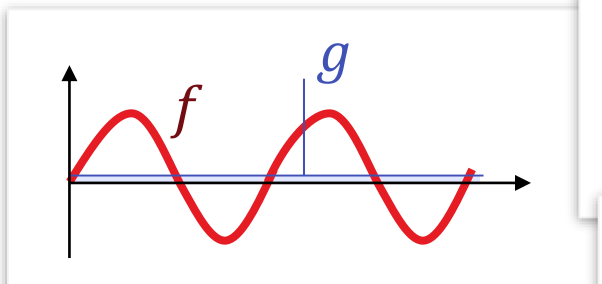
Example

- $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5)$
- $\mathbf{x} \mapsto x_4$ is a linear maps
- Represented by $\langle (0,0,0,1,0), \mathbf{x} \rangle$
- "Linear form": $\mathbf{x} \mapsto \langle (0,0,0,1,0), \mathbf{x} \rangle$, in short, $\langle \cdot, (0,0,0,1,0) \rangle$, shorter: $(0,0,0,1,0) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}^{\mathrm{T}}$

Linear Forms in Function Spaces

In function spaces

- Picking entries by x-axis is a linear operation
- *Cannot* be represented by projection to another function (multiplication with "co-vector")



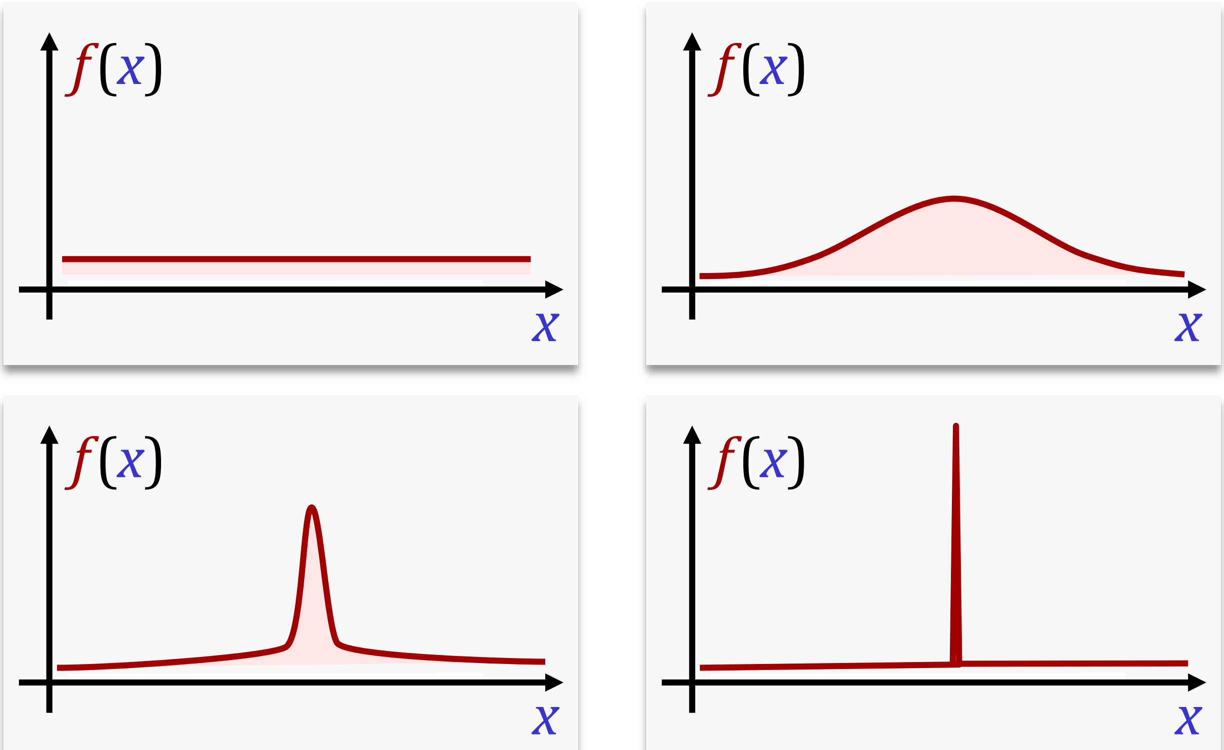
$$g(x) = \begin{cases} 1, \text{if } x = 4 \\ 0, \text{elsewhere} \end{cases}$$

$$\int_{\mathbb{R}} f(x)g(x)dx = 0$$

Example

- $f: \mathbb{R} \to \mathbb{R}, \quad f(x) = \sin(x)$
- $L: (\mathbb{R} \to \mathbb{R}) \to \mathbb{R}, \quad L: f \mapsto f(4.0)$ is a linear map
- A function g with $\langle g, f \rangle = f(4.0)$ does not exist

Dirac's "Delta Function"



$$\int_{\Omega} f(x)dx = 1$$

Dirac Delta "Function"

- $\int_{\mathbb{R}} \delta(x)dx = 1$, zero everywhere but at $x = 0$
- Idealization ("distribution") – think of very sharp peak

Distributions

Distributions

- Adding all linear forms to the vector space
  - All linear mappings from the function space to $\mathbb{R}$
- This makes the situation symmetric
- δ is a distribution, not a (traditional) function

Formalization

- Different approaches (details beyond our course)
  - Limits of “bumps”
  - Space of linear-forms (“co-vectors”, “dual functions”)
  - Difference of complex functions on Riemann sphere (exotic)