Linear transformation of a vector $x$. The ideas presented here are quite general. They go beyond the traditional matrix-vector type seen in linear algebra. Furthermore, they do not deal with basis and are equally valid for any set of basis. While some people consider transformation from one dimension to another (say, $\mathbb{R}^m \rightarrow \mathbb{R}^n$) or from one type of vector to another type (say, from a column of numbers to a row of numbers), or include complex numbers, we shall restrict ourselves to the same type of vectors in the same n-dimensional real space.

Instructor: Nam Sun Wang

**Transformation** in LVS is a rule that maps a vector in the LVS to another vector in the same LVS. An analogy is an ordinary function that takes in a scalar number and assigns another scalar number. A transformation is similar to a scalar function, except it takes in a vector and assigns another vector. In essence, a transformation is a vector function of vectors. It is also commonly referred to as a tensor or an operator. When we speak of transformation, we need not restrict ourselves to just real Euclidean space nor a LVS of finite dimension. We can make up any rules to map a vector to another vector, but some rules are more useful than others. There are several generally accepted notations; below are some examples, and $x$ and $y$ are vectors.

$A \cdot x = y \quad A \cdot x = y \quad f(x) = y \quad \mathcal{L}(x) = y$

Note that the first notation above does not necessarily mean $A$ is a matrix (which is true only for the special case when the vector is a column of numbers or a row of numbers and the matrix $A$ must be a square one; in the latter case or a row of numbers, watch out how we define the operation to preserve a row of numbers). There is no requirement that we must use a different font to denote transformation to reduce confusion. Furthermore, a transformation has nothing to do with basis. With the $A \cdot x = y$ notation, we do not imply the transformation $A$ is a linear one either.

**Linear Transformation.** A transformation $A$ is a linear transformation if it satisfies the following two properties on 1) addition of two vectors and 2) multiplication of a vector by a scalar. These properties must hold for every vector in LVS, not just some vectors.

1. Distributive $A \cdot (x + y) = A \cdot x + A \cdot y$
2. Associative $A \cdot (\alpha \cdot x) = (\alpha \cdot A \cdot x)$

- **Example:** the "traditional," strictly linear scalar functions of scalars (i.e., a 1st degree polynomial of the following form without the constant term) is a linear transformation.

$$f(x) = \alpha \cdot x$$

How about the more general "linear" scalar functions of scalars (i.e., a 1st degree polynomial of the following form with a constant term $\alpha$)?

$$f(x) = \alpha + \beta \cdot x$$

- **Example:**

$$A \cdot x = y \quad \text{where} \quad x \text{ and } y \text{ are two columns of real numbers. The rule is defined as:}$$

$$y_i = \sum_{j=1}^{n} a_{ij} \cdot x_j \quad \text{for} \quad i = 1, 2 \ldots n$$

- **Example:**

Rotation of arrows. $R \uparrow = \longrightarrow$ or $\mathcal{R}(\uparrow) = \longrightarrow$
• Example: Differentiation, What is the subject of this operator? In other words, this differentiation operator operates on what kinds of vectors? Arrows? Columns of numbers? Continuous functions with 1st derivatives?

\[ \frac{d}{dt} f(t) \]

\[ \frac{d}{dt} \left( \frac{d}{dt} f(t) \right) = \frac{d^2}{dt^2} f(t) \]

\( \mathcal{D}(f) \) \text{ ... another notation for a differential operator.}

We can show that the transformation \( \mathcal{D} \) satisfies the two linear properties. Thus, \( \mathcal{D} \) is a linear transformation.

1. Distributive  
\[ \mathcal{D}(f + g) = \frac{d}{dt}(f(t) + g(t)) = \frac{d}{dt} f(t) + \frac{d}{dt} g(t) = \mathcal{D} f + \mathcal{D} g \]

2. Associative  
\[ \mathcal{D}(\alpha f) = \frac{d}{dt}(\alpha f(t)) = \alpha \frac{d}{dt} f(t) = \alpha (\mathcal{D} f) \]

• Example: A more general differential operator. Here is a second-order ODE operator.

\[ L = \frac{a(t) \frac{d^2}{dt^2} + b(t) \frac{d}{dt} + c(t)}{dt} \cdot f(t) = \frac{a(t) \frac{d^2}{dt^2} f(t) + b(t) \frac{d}{dt} f(t) + c(t) f(t)}{dt} \]

ODE(\( f \)) \text{ ... another notation for a differential operator.}

Bessel differential operator:  
\[ ODE(y) = \left( t^2 \frac{d^2}{dt^2} + t \frac{d}{dt} + t^2 \right) y = \frac{d}{dt} \left( t^2 \frac{d}{dt} + t \frac{d}{dt} + t^2 \right) y \]

We can show that the transformation ODE satisfies the two linear properties. Thus, ODE is a linear transformation.

1. Distributive  
\[ ODE(f + g) = \left( t^2 \frac{d^2}{dt^2} + t \frac{d}{dt} + t^2 \right) (f + g) = \left( t^2 \frac{d^2}{dt^2} + t \frac{d}{dt} + t^2 \right) f + \left( t^2 \frac{d^2}{dt^2} + t \frac{d}{dt} + t^2 \right) g \]

\[ = ODEf + ODEg \]

2. Associative  
\[ ODE(\alpha f) = \left( t^2 \frac{d^2}{dt^2} + t \frac{d}{dt} + t^2 \right) (\alpha f) = \alpha \left( t^2 \frac{d^2}{dt^2} + t \frac{d}{dt} + t^2 \right) (f(t)) = \alpha (ODE(f)) \]

• Example: Let the LVS be the space of all polynomials of degree \( n \) or less.

\[ P_n(t) = a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n \]

Define a linear transformation \( \mathcal{D} \) such that  
\[ \mathcal{D} P = \frac{d}{dt} f(t) \]

We can show that \( \mathcal{D} \) satisfies the two linear properties.

• Non-Example: What if the LVS is composed of all real functions? Answer: No, because the derivative may not exist. What if we tighten the membership requirement of the LVS so that the LVS is composed of all real, continuous functions that possess the 1st derivatives? Answer: No, because the resulting function may get kicked out from the LVS for not having a 1st derivative.
Example: Integration. We can show that the integral operator is a linear one. (What is the LVS in which the integral operator is defined in? In other words, what are the membership requirements?)

\[ \mathcal{L} f = \int_{0}^{t} f(\tau) \, d\tau \]

\[ \mathcal{L}(f) \quad \text{... another notation for an integral operator.} \]

What if the LVS is the space of all polynomials of degree \( n \) or less? Answer: No, \( \mathcal{L} \) is not a linear transformation for this LVS, because the resulting function is a polynomial of degree \( n+1 \), which is excluded from the membership.

\[ p_n(t) = a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n \]

\[ \mathcal{L} p_n = \int_{0}^{t} p_n(\tau) \, d\tau = a_0 + a_1 \frac{t^2}{2} + a_2 \frac{t^3}{3} + \ldots + a_n \frac{t^{n+1}}{n+1} \]

What if the LVS is the space of all polynomials of degree \( n \) or less, as before, but now with a slightly different integral operator, like the one defined below? Answer: Yes, because the result of the linear transformation now lies within the same LVS.

Define \( \mathcal{L} p_n = \frac{1}{t} \int_{0}^{t} p_n(\tau) \, d\tau \)

\[ \mathcal{L} p_n = a_0 + a_1 \frac{t^2}{2} + a_2 \frac{t^3}{3} + \ldots + a_n \frac{t^n}{n+1} \]

Example: Convolution integral.

\[ \mathcal{L} f = \int_{0}^{1} f(\tau) \cdot g(t-\tau) \, d\tau \quad \text{or} \quad \mathcal{L} f = \int_{0}^{1} f(\tau) \cdot g(t+\tau) \, d\tau \]

Example: Laplace Transform.

\[ \mathcal{L} f = \int_{0}^{\infty} e^{-st} \cdot f(t) \, dt = F(s) \quad \mathcal{L}(f(t)) = F(s) \quad \text{... another common notation} \]

Example: Fourier Sine Transform.

\[ \mathcal{F} f = \frac{1}{\sqrt{\pi}} \int_{0}^{\infty} \sin(\omega t) \cdot f(t) \, dt = F(\omega) \quad \omega \geq 0 \quad f(t) = e^{-t} \int_{0}^{\omega} \sin(\omega t) \cdot e^{-t} \, dt = \frac{\omega}{1 + \omega^2} \]

Example: Fourier Cosine Transform.

\[ \mathcal{F} f = \frac{1}{\sqrt{\pi}} \int_{0}^{\infty} \cos(\omega t) \cdot f(t) \, dt = F(\omega) \quad \omega \geq 0 \int_{0}^{\infty} \cos(\omega t) \cdot e^{-t} \, dt = \frac{1}{1 + \omega^2} \]

Example: Fourier Transform. \( i \quad \text{... imaginary number} \)

\[ \mathcal{F} f = \int_{-\infty}^{\infty} e^{i2\pi\omega t} \cdot f(t) \, dt = F(\omega) \quad \mathcal{F}(f) = F(\omega) \quad \text{... another common notation} \]
Example: Projector operator. Consider any real Euclidean space \( \mathbb{R} \). Let \( \mathcal{P} \) be a projector that projects a vector \( x \) in any Euclidean space onto a finite dimensional subspace of \( \mathbb{R} \).

Note that this is just a notation. It does not imply multiplication of \( x \) by an operator \( \mathcal{P} \), which makes no sense.

The usual algebraic operations between two linear transformations (equality, addition of two transformations, multiplication of a transformation by a scalar, multiplication of two transformations) are defined in terms of the result of the transformation, never just the transformations. Remember, the result of transformation is another vector. Strictly speaking, transformations all by themselves has no meaning. Transformations are not objects/things/animals like the way vectors are.

Two linear transformations \( \mathcal{A} \) and \( \mathcal{B} \) are equal if for every vector \( x \) in the LVS, we have, \( \mathcal{A} \cdot x = \mathcal{B} \cdot x \)

Thus, when we speak of two equal linear transformations, we are really referring to the two equal vectors that result from applying the transformations to the same vector.

The sum/addition of two linear transformations \( \mathcal{A} \) and \( \mathcal{B} \) is defined as

\[
\mathcal{C} = \mathcal{A} + \mathcal{B}
\]

if for every vector \( x \) in the LVS, we have,

\[
\mathcal{C} \cdot x = \mathcal{A} \cdot x + \mathcal{B} \cdot x
\]

The product/multiplication of a transformation \( \mathcal{A} \) by a scalar \( \alpha \) is defined, if for every vector \( x \) in the LVS, we have

\[
(\alpha \cdot \mathcal{A}) \cdot x = \mathcal{A} \cdot (\alpha \cdot x)
\]

The product/multiplication of two linear transformations \( \mathcal{A} \) and \( \mathcal{B} \) is defined as

\[
\mathcal{C} = \mathcal{A} \cdot \mathcal{B}
\]

if for every vector \( x \) in the LVS, we have,

\[
\mathcal{C} \cdot x = \mathcal{A} \cdot (\mathcal{B} \cdot x)
\]

In general, tensor products are not commutative: \( \mathcal{A} \circ \mathcal{B} \neq \mathcal{B} \circ \mathcal{A} \)

We can demonstrate \( \mathcal{A} \circ \mathcal{B} \neq \mathcal{B} \circ \mathcal{A} \) with matrices. We can also demonstrate this fact with any one of the examples of linear transformations mentioned above.

- Example. Change one column of numbers to another column of numbers (i.e., the usual matrix). We can easily show \( \mathcal{A} \circ \mathcal{B} \neq \mathcal{B} \circ \mathcal{A} \) to be true..

- Example. All polynomials of degree \( n \) or less

Define two linear transformations \( \mathcal{A} \) and \( \mathcal{B} \)

\[
\mathcal{A} \cdot P \overset{d}{\underset{d}{=}} \frac{d}{dt} P(t)
\]

\[
\mathcal{B} \cdot P \overset{\int}{\underset{\int}{=}} \int_0^t P(\tau) \, d\tau
\]

Apply two linear transformations in two different sequences

\[
\mathcal{A} \cdot (\mathcal{B} \cdot P) \overset{d}{\underset{d}{=}} \frac{d}{dt} \left( \int_0^t P(\tau) \, d\tau \right) = P(t)
\]

\[
\mathcal{B} \cdot (\mathcal{A} \cdot P) \overset{\int}{\underset{\int}{=}} \int_0^t \left( \frac{d}{d\tau} P(\tau) \right) \, d\tau = \int_0^t P(\tau) \, d\tau - P(0)
\]

\[
\mathcal{B} \cdot (\mathcal{A} \cdot P) - \mathcal{A} \cdot (\mathcal{B} \cdot P) = (P(t) - P(0)) - P(t) = P(0) = 0 \quad \rightarrow \quad \mathcal{A} \cdot (\mathcal{B} \cdot P) = \mathcal{B} \cdot (\mathcal{A} \cdot P) \quad \rightarrow \quad \mathcal{A} \cdot \mathcal{B} = \mathcal{B} \cdot \mathcal{A}
\]
Example. All polynomials of degree n or less

Define two linear transformations $A$ and $B$

$$A\ P = \frac{d}{dt} P(t)$$
$$B\ P = \frac{1}{t} \int_0^t P(\tau) \, d\tau$$

Apply two linear transformations in two different sequences

$$A(B\ P) = \frac{d}{dt} \left( \frac{1}{t} \int_0^t P(\tau) \, d\tau \right) = \frac{1}{t} P(t) - \frac{1}{t^2} \int_0^t P(\tau) \, d\tau$$
$$B(A\ P) = \frac{1}{t} \int_0^t \frac{d}{d\tau} P(\tau) \, d\tau = \frac{1}{t} \int_0^t 1 \, d\tau = \frac{0}{t}$$

$$B(A\ P) - A(B\ P) = \frac{1}{t^2} \int_0^t (P(t) - P(0)) \frac{0}{t} = 0 \quad \rightarrow \quad A(B\ P) = B(A\ P) \quad \rightarrow \quad A \cdot B = B \cdot A$$
Matrix to characterize linear transformation in n-dimensional LVS. If we know how a linear transformation acts on a set of basis vectors, we can easily find how it acts on any vectors (obviously they must belong to the same LVS). In a n-dimensional LVS, a linear transformation turns a set of n linearly independent basis vectors \{x_1, x_2, ..., x_n\} into a new set of n vectors \{y_1, y_2, ..., y_n\} (which, again, must continue to belong to the same LVS). These output vectors from the linear transformation, in turn, can be expressed as a linear combination of the basis vectors \{x_1, x_2, ..., x_n\}. Any linear transformation \(A\) in n-dimensional LVS can be expressed as a matrix \(A_{\text{matrix}}\). (Note that the last statement is true for a finite dimensional LVS; it is not true for LVS of an infinity dimension, e.g., all continuous functions, because we need an infinite number of basis to describe a random continuous function.)

Linear transformation on the basis vectors (This formulation does not work for \(\infty\)-dimension.)

\[
\begin{align*}
A \cdot x_1 &= y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\
A \cdot x_2 &= y_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\
\vdots \\
A \cdot x_n &= y_n = a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n
\end{align*}
\]

In a compact matrix notation (some people abhor this, because we are mixing vectors and columns|rows of vectors & also vectors that are rectangular-wise arranged set of numbers and matrices). Thus, the action of a linear transformation can be characterized by a matrix \(A_{\text{matrix}}\):

Compact form with vectors \(x\) & \(y\) listed in a column

\[
\begin{bmatrix}
A \cdot x_1 \\
A \cdot x_2 \\
\vdots \\
A \cdot x_n
\end{bmatrix} = \begin{bmatrix} y_1 \\
y_2 \\
\vdots \\
y_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\
a_{21} & a_{22} & \ldots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\
x_2 \\
\vdots \\
x_n \end{bmatrix}
\]

Compact form with vectors \(x\) & \(y\) listed in a row

\[
\begin{bmatrix} A \cdot x_1 & A \cdot x_2 & \ldots & A \cdot x_n \end{bmatrix} = \begin{bmatrix} y_1 & y_2 & \ldots & y_n \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix}^T \begin{bmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\
a_{12} & a_{22} & \ldots & a_{n2} \\
\vdots & \vdots & \ddots & \vdots \\
a_{1n} & a_{2n} & \ldots & a_{nn} \end{bmatrix}
\]

Here, "\(A\)" symbolizes linear transformation; whereas, "\(A_{\text{matrix}}\)" is a matrix that characterizes the linear transformation.
Any vector \( \mathbf{x} \) in the LVS can be represented as a linear combination of the \( n \) basis vectors \( \{ \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \} \):

\[
\mathbf{x} = \xi_1 \mathbf{x}_1 + \xi_2 \mathbf{x}_2 + \ldots + \xi_n \mathbf{x}_n
\]

The action of the linear transformation on an arbitrarily chosen vector \( \mathbf{x} \) is described as:

\[
y = \mathbf{A} \mathbf{x} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}
\]

compact matrix notation with vectors \( \mathbf{x} \) & \( \mathbf{y} \) listed in a column (again, remember, some people abhor this compact notation because we are mixing a column|row of vectors with a column|row of scalars)

\[
y = \mathbf{A} \mathbf{x} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}
\]

compact matrix notation with vectors \( \mathbf{x} \) & \( \mathbf{y} \) listed in a row

The choice of basis is not unique. The same LVS can be expressed based on another set of basis \( \{ \mathbf{x}'_1, \mathbf{x}'_2, \ldots, \mathbf{x}'_n \} \), which is related to the first set as:

\[
\mathbf{x}' = \mathbf{T} \mathbf{x} \quad \mathbf{x}' = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}
\]

Linear transformation on the basis vectors (This formulation does not work for \( \infty \)-dimension.)

\[
\mathbf{A} \mathbf{x}'_1 = \mathbf{a}'_1 \mathbf{x}'_1 + \mathbf{a}'_2 \mathbf{x}'_2 + \ldots + \mathbf{a}'_n \mathbf{x}'_n = \sum_{i} a'_{i} \mathbf{x}'_i
\]

\[
\mathbf{A} \mathbf{x}'_2 = \mathbf{a}'_2 \mathbf{x}'_1 + \mathbf{a}'_2 \mathbf{x}'_2 + \ldots + \mathbf{a}'_n \mathbf{x}'_n = \sum_{i} a'_{2i} \mathbf{x}'_i
\]

\[
\mathbf{A} \mathbf{x}'_n = \mathbf{a}'_n \mathbf{x}'_1 + \mathbf{a}'_n \mathbf{x}'_2 + \ldots + \mathbf{a}'_{n} \mathbf{x}'_n = \sum_{i} a'_{ni} \mathbf{x}'_i
\]
In a compact matrix notation with vectors \( x \) & \( y \) listed in a column

\[
\begin{pmatrix}
    A'x_1 & y_1 \\
    A'x_2 & y_2 \\
    \vdots & \vdots \\
    A'x_n & y_n
\end{pmatrix} =
\begin{pmatrix}
    a'_{11} & a'_{12} & \ldots & a'_{1n} \\
    a'_{21} & a'_{22} & \ldots & a'_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    a'_{n1} & a'_{n2} & \ldots & a'_{nn}
\end{pmatrix}
\begin{pmatrix}
    a_{11} & a_{12} & \ldots & a_{1n} \\
    a_{21} & a_{22} & \ldots & a_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{n1} & a_{n2} & \ldots & a_{nn}
\end{pmatrix}
\begin{pmatrix}
    x_1 \\
    x_2 \\
    \vdots \\
    x_n
\end{pmatrix}
\]

Thus,

\[
\begin{pmatrix}
    [A'x_1] \\
    [A'x_2] \\
    \vdots \\
    [A'x_n]
\end{pmatrix} = \begin{pmatrix}
    t_{11} & t_{12} & \ldots & t_{1n} \\
    t_{21} & t_{22} & \ldots & t_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    t_{n1} & t_{n2} & \ldots & t_{nn}
\end{pmatrix}
\begin{pmatrix}
    a'_{11} & a'_{12} & \ldots & a'_{1n} \\
    a'_{21} & a'_{22} & \ldots & a'_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    a'_{n1} & a'_{n2} & \ldots & a'_{nn}
\end{pmatrix}
\begin{pmatrix}
    x_1 \\
    x_2 \\
    \vdots \\
    x_n
\end{pmatrix}
\]

\( A' \) matrix \( T \) matrix = \( T \) matrix \( \cdot \) \( A \) matrix

If \( x_i \) & \( x'_i \) are two orthonormal sets, then \( T \) matrix is orthogonal.

In a compact matrix notation with vectors \( x \) & \( y \) listed in a row

\[
\begin{pmatrix}
    A'x_1 & A'x_2 & \ldots & A'x_n
\end{pmatrix} =
\begin{pmatrix}
    a'_{11} & a'_{12} & \ldots & a'_{1n} \\
    a'_{21} & a'_{22} & \ldots & a'_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    a'_{n1} & a'_{n2} & \ldots & a'_{nn}
\end{pmatrix}
\begin{pmatrix}
    a_{11} & a_{12} & \ldots & a_{1n} \\
    a_{21} & a_{22} & \ldots & a_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{n1} & a_{n2} & \ldots & a_{nn}
\end{pmatrix}
\begin{pmatrix}
    y_1 \\
    y_2 \\
    \vdots \\
    y_n
\end{pmatrix}
\]

Thus,

\[
\begin{pmatrix}
    [A'x_1] \\
    [A'x_2] \\
    \vdots \\
    [A'x_n]
\end{pmatrix} = \begin{pmatrix}
    t_{11} & t_{12} & \ldots & t_{1n} \\
    t_{21} & t_{22} & \ldots & t_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    t_{n1} & t_{n2} & \ldots & t_{nn}
\end{pmatrix}
\begin{pmatrix}
    a'_{11} & a'_{12} & \ldots & a'_{1n} \\
    a'_{21} & a'_{22} & \ldots & a'_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    a'_{n1} & a'_{n2} & \ldots & a'_{nn}
\end{pmatrix}
\begin{pmatrix}
    x_1 \\
    x_2 \\
    \vdots \\
    x_n
\end{pmatrix}
\]

\( A' \) matrix \( T \) matrix = \( T \) matrix \( \cdot \) \( A \) matrix
Example. Derivative of polynomials (basis=power series). All polynomials of degree \(n\) or less \(P(t) = a_0 + a_1 t + \ldots + a_n t^n\) produce another set of polynomials. If the LVS is all polynomials of any degree, the derivative operator acting on these power series produce another set of polynomials. If the LVS is all polynomials of any degree, a finite set of basis functions exist and a compact notation does not exist.

Compact matrix notation with vectors \(x\) & \(y\) listed in a column

\[
\begin{bmatrix}
  f_0 \\
  f_1 \\
  f_2 \\
  \vdots \\
  f_n
\end{bmatrix} = \begin{bmatrix}
  0 & 0 & 0 & \ldots & 0 \\
  0 & 1 & 0 & \ldots & 0 \\
  0 & 0 & 2 & \ldots & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & 0 & \ldots & n
\end{bmatrix} \begin{bmatrix}
  f_0 \\
  f_1 \\
  f_2 \\
  \vdots \\
  f_n
\end{bmatrix}
\]

Compact matrix notation with functions \(f\) listed in a row

\[
\begin{bmatrix}
  f_0 \\
  f_1 \\
  f_2 \\
  \vdots \\
  f_n
\end{bmatrix}^T \begin{bmatrix}
  f_0 \\
  f_1 \\
  f_2 \\
  \vdots \\
  f_n
\end{bmatrix} = \begin{bmatrix}
  0 & 1 & 0 & \ldots & 0 \\
  0 & 0 & 2 & \ldots & 0 \\
  0 & 0 & 0 & \ldots & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & 0 & \ldots & 0
\end{bmatrix}
\]

Any function/\(\xi\) vector in the LVS (that is, any polynomials of degree \(n\) or less) can be represented as a linear combination of the \(n+1\) basis functions

\[
y = \xi_0 f_0 + \xi_1 f_1 + \ldots + \xi_n f_n
\]

generally accepted

\[
\mathcal{D} y = \mathcal{D} \begin{bmatrix}
  f_0 \\
  f_1 \\
  f_2 \\
  \vdots \\
  f_n
\end{bmatrix} = \begin{bmatrix}
  \xi_0 f_0 + \xi_1 f_1 + \ldots + \xi_n f_n \\
  \xi_1 f_0 + \xi_2 f_1 + \ldots + \xi_n f_n \\
  \xi_2 f_0 + \xi_3 f_1 + \ldots + \xi_n f_n \\
  \vdots \\
  \xi_n f_0 + \xi_1 f_1 + \ldots + \xi_n f_n
\end{bmatrix}
\]

Compact matrix notation with vectors \(f\) in a column

\[
\begin{bmatrix}
  \xi_0 \\
  \xi_1 \\
  \xi_2 \\
  \vdots \\
  \xi_n
\end{bmatrix} \begin{bmatrix}
  f_0 \\
  f_1 \\
  f_2 \\
  \vdots \\
  f_n
\end{bmatrix} = \begin{bmatrix}
  f_0 \\
  f_1 \\
  f_2 \\
  \vdots \\
  f_n
\end{bmatrix} \mathcal{D} \begin{bmatrix}
  \xi_0 \\
  \xi_1 \\
  \xi_2 \\
  \vdots \\
  \xi_n
\end{bmatrix}
\]

Compact matrix notation with vectors \(f\) in a row

\[
\begin{bmatrix}
  f_0 \\
  f_1 \\
  f_2 \\
  \vdots \\
  f_n
\end{bmatrix}^T \mathcal{D} \begin{bmatrix}
  \xi_0 \\
  \xi_1 \\
  \xi_2 \\
  \vdots \\
  \xi_n
\end{bmatrix} = \begin{bmatrix}
  \xi_0 f_0 + \xi_1 f_1 + \ldots + \xi_n f_n
\end{bmatrix}^T \mathcal{D}
\]
\[ \mathcal{D} \mathbf{y} = \begin{bmatrix} f_0 & f_1 & \ldots & f_n \end{bmatrix} \begin{bmatrix} \xi_0 & 1 & \ldots & \xi_n \end{bmatrix} = \begin{bmatrix} f_{\text{prime}}^T \cdot \xi \end{bmatrix} = f_{\text{prime}}^T \cdot \xi \]

**Example.** Derivative of polynomials (basis=Legendre polynomials). All polynomials of degree \( n \) or less. We can use Legendre polynomials \( P \) as the \( n+1 \) basis functions/vectors. We find the derivative operator acting on these power series produce another set of polynomials. Note that the choice of a set of basis vectors is not unique. The representation of a given vector depends on the choice of basis vectors (see change-of-basis formula). The representation of the action of a linear transformation as a matrix depends on the choice of basis vectors and is not unique.

Compact matrix notation with vectors \( f \) & \( P \) in a column

\[
\begin{bmatrix}
P_0 \\
P_1 \\
\vdots \\
P_n
\end{bmatrix} = 
\begin{bmatrix}
1 & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1
\end{bmatrix} 
\begin{bmatrix}
f_0 \\
f_1 \\
\vdots \\
f_n
\end{bmatrix}
\]

P(x) = A \cdot f(x) \quad P'(x) = A \cdot \text{prime} \cdot A^{-1} \cdot P(x) = \text{Prime} \cdot P(x)

where \quad \text{Prime} = A \cdot \text{prime} \cdot A^{-1}

Compact matrix notation with vectors \( f \) & \( P \) in a row

\[
\begin{bmatrix}
P_0 & P_1 & \ldots & P_n
\end{bmatrix} = 
\begin{bmatrix}
f_0 & f_1 & \ldots & f_n
\end{bmatrix} \cdot A^T 
\]

\[
\begin{bmatrix}
P(x) = f(x) \cdot A^T \\
P'(x) = \text{prime}^T \cdot A^T = P(x) \cdot (A^T)^{-1} \cdot \text{prime}^T \cdot A^T
\end{bmatrix}
\]

Thus, the matrix that characterizes the action of derivative on Legendre basis functions is

\[
\text{Prime} = A \cdot \text{prime} \cdot A^{-1}
\]

Note that the same derivative operator is characterized by different matrices, depending on the basis functions/vectors.

**Example.** Integration of polynomials. All polynomials \( P \) of degree \( n \) or less. We can use power series \( \{1, x, x^2, \ldots, x^n\} \) as the \( n+1 \) basis functions/vectors. The integration operator acting on these power series produce another set of polynomials.

\[
\mathcal{B} \mathbf{P} = \int_0^t P(\tau) \, d\tau
\]

Any members in the LVS (that is, any polynomials of degree \( n \) or less) can be represented as a linear combination of the \( n+1 \) basis functions

\[
P = \xi_0 \cdot f_0 + \xi_1 \cdot f_1 + \ldots + \xi_n \cdot f_n
\]

generally accepted

\[
\mathcal{B} \mathbf{P} = \begin{bmatrix}
\xi_0 & f_0 \\
\xi_1 & f_1 \\
\vdots & \vdots
\end{bmatrix}
\]

Compact matrix notation with vectors \( f \) in a column

\[
\begin{bmatrix}
f_0 \\
f_1 \\
\vdots
\end{bmatrix} = 
\begin{bmatrix}
1 & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1
\end{bmatrix} 
\begin{bmatrix}
f_0 = 1 \\
f_1 = x
\end{bmatrix}
\]
compact matrix notation with vectors $f$ in a row

\[ \begin{bmatrix} f_0 & f_1 & f_2 & \cdots & f_n \end{bmatrix} \]

- **Example Projection** of a given vector $y$ in $n$-dimensional space onto basis vector $x_2$.

\[ P_2 y = \begin{bmatrix} y \cdot x_2 \\ x_2 \cdot x_2 \end{bmatrix} - x_2 \] projection of $y$ along $x_2$

Action of $P_2$ on each of the basis vectors

- action of $P_2$ on $x_1$  
  \[ P_2 x_1 = y \cdot \frac{x_1 \cdot x_2}{x_2 \cdot x_2} \]

- action of $P_2$ on $x_2$  
  \[ P_2 x_2 = y \cdot \frac{x_2 \cdot x_2}{x_2 \cdot x_2} = 1 \cdot x_2 \]

- action of $P_2$ on $x_n$  
  \[ P_2 x_n = y \cdot \frac{x_n \cdot x_2}{x_2 \cdot x_2} \]

compact matrix notation with vectors $x$ & $y_{\text{proj}}$ in a column

\[ P_2 x = \begin{bmatrix} P_2 x_1 \\ P_2 x_2 \\ \vdots \\ P_2 x_n \end{bmatrix} = \begin{bmatrix} y \cdot x_{\text{proj2}} x_1 \\ y \cdot x_{\text{proj2}} x_2 \\ \vdots \\ y \cdot x_{\text{proj2}} x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \frac{x_1 x_2}{x_2 x_2} \cdots x_n \]

where only one column of $A_{\text{matrix}}$ is nonzero

compact matrix notation with vectors $x$ & $y_{\text{proj}}$ in a row

\[ P_2 x = \begin{bmatrix} P_2 x_1 & P_2 x_2 & \cdots & P_2 x_n \end{bmatrix} = \begin{bmatrix} y \cdot x_{\text{proj2}} x_1 & y \cdot x_{\text{proj2}} x_2 & \cdots & y \cdot x_{\text{proj2}} x_n \end{bmatrix} \]

\[ = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \frac{x_1 x_2}{x_2 x_2} & \frac{x_2 x_2}{x_2 x_2} & \cdots & \frac{x_n x_2}{x_2 x_2} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \]
Any vector \( y \) in the LVS can be represented as a linear combination of the \( n \) basis vectors \( \{x_1, x_2, \ldots, x_n\} \)
\[
y = \xi_1 x_1 + \xi_2 x_2 + \ldots + \xi_n x_n
\]
The action of the linear transformation on an arbitrarily chosen vector \( y \) is described as
\[
y \overset{P_2}{\rightarrow} \xi_1 y_{proj2,x1} + \xi_2 y_{proj2,x2} + \ldots + \xi_n y_{proj2,xn}
\]
compact matrix notation with vectors \( y \) and \( y_{proj} \) listed in a column (again, remember, some people abhor this compact notation because we are mixing a column|row of vectors with a column|row of scalars)

\[
\begin{bmatrix}
y_{proj2,x1} \\
y_{proj2,x2} \\
\vdots \\
y_{proj2,xn}
\end{bmatrix} =
\begin{bmatrix}
x_1, x_2 \ldots x_n
\end{bmatrix}^T 
\begin{bmatrix}
0 & \frac{x_1, x_2}{x_2, x_2} & \ldots & 0 \\
0 & \frac{x_2, x_2}{x_2, x_2} & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & \frac{x_n, x_2}{x_2, x_2} & \ldots & 0
\end{bmatrix} 
\begin{bmatrix}
\xi_1 \\
\xi_2 \\
\vdots \\
\xi_n
\end{bmatrix}
\]
compact matrix notation with vectors \( y \) & \( y_{proj} \) listed in a row

\[
\begin{bmatrix}
y_{proj2,x1} & y_{proj2,x2} & \ldots & y_{proj2,xn}
\end{bmatrix} =
\begin{bmatrix}
x_1, x_2 \ldots x_n
\end{bmatrix} 
\begin{bmatrix}
0 & 0 & \ldots & 0 \\
\frac{x_1, x_2}{x_2, x_2} & \frac{x_2, x_2}{x_2, x_2} & \ldots & \frac{x_n, x_2}{x_2, x_2} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 0
\end{bmatrix} 
\begin{bmatrix}
\xi_1 \\
\xi_2 \\
\vdots \\
\xi_n
\end{bmatrix}
\]
If the basis are orthonormal, then

\[
A_{matrix} =
\begin{bmatrix}
0 & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0
\end{bmatrix}
\]
• **Example Projection** of a given vector \( y \) in \( n \)-dimensional space onto \( n \) basis vectors \( \{x_1, x_2, \ldots, x_n\} \).

Projection of \( y \) along \( x_1 \):

\[
P_1y = \left( \frac{y \cdot x_1}{x_1 \cdot x_1} \right) x_1 = \left( \frac{\xi_1}{\frac{x_1 \cdot x_1}{x_1 \cdot x_1}} \right) x_1 + \left( \frac{\xi_2}{\frac{x_1 \cdot x_1}{x_1 \cdot x_1}} \right) x_2 + \ldots + \left( \frac{\xi_n}{\frac{x_1 \cdot x_1}{x_1 \cdot x_1}} \right) x_n \]

Projection of \( y \) along \( x_2 \):

\[
P_2y = \left( \frac{y \cdot x_2}{x_2 \cdot x_2} \right) x_2 = \left( \frac{\xi_1}{\frac{x_2 \cdot x_2}{x_2 \cdot x_2}} \right) x_1 + \left( \frac{\xi_2}{\frac{x_2 \cdot x_2}{x_2 \cdot x_2}} \right) x_2 + \ldots + \left( \frac{\xi_n}{\frac{x_2 \cdot x_2}{x_2 \cdot x_2}} \right) x_n \]

...  

Projection of \( y \) along \( x_n \):

\[
P_ny = \left( \frac{y \cdot x_n}{x_n \cdot x_n} \right) x_n = \left( \frac{\xi_1}{\frac{x_n \cdot x_n}{x_n \cdot x_n}} \right) x_1 + \left( \frac{\xi_2}{\frac{x_n \cdot x_n}{x_n \cdot x_n}} \right) x_2 + \ldots + \left( \frac{\xi_n}{\frac{x_n \cdot x_n}{x_n \cdot x_n}} \right) x_n \]

Combining the above \( n \) projection operators in a column:

\[
\begin{bmatrix}
P_1y \\
P_2y \\
\vdots \\
P_ny \\
\end{bmatrix} =
\begin{bmatrix}
\left( \frac{y \cdot x_1}{x_1 \cdot x_1} \right) & 0 & \cdots & 0 \\
0 & \left( \frac{y \cdot x_2}{x_2 \cdot x_2} \right) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \left( \frac{y \cdot x_n}{x_n \cdot x_n} \right) \\
\end{bmatrix}
\begin{bmatrix}
\xi_1 \\
\xi_2 \\
\xi_n \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
P_1 \cdot x_1 & P_1 \cdot x_2 & \cdots & P_1 \cdot x_n \\
P_2 \cdot x_1 & P_2 \cdot x_2 & \cdots & P_2 \cdot x_n \\
\vdots & \vdots & \ddots & \vdots \\
P_n \cdot x_1 & P_n \cdot x_2 & \cdots & P_n \cdot x_n \\
\end{bmatrix}
\begin{bmatrix}
\xi_1 \\
\xi_2 \\
\xi_n \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
\frac{x_1 \cdot x_1}{x_1 \cdot x_1} & \frac{x_2 \cdot x_1}{x_1 \cdot x_1} & \cdots & \frac{x_n \cdot x_1}{x_1 \cdot x_1} \\
\frac{x_1 \cdot x_2}{x_1 \cdot x_2} & \frac{x_2 \cdot x_2}{x_1 \cdot x_2} & \cdots & \frac{x_n \cdot x_2}{x_1 \cdot x_2} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{x_1 \cdot x_n}{x_1 \cdot x_n} & \frac{x_2 \cdot x_n}{x_1 \cdot x_n} & \cdots & \frac{x_n \cdot x_n}{x_1 \cdot x_n} \\
\end{bmatrix}
\begin{bmatrix}
\xi_1 \\
\xi_2 \\
\xi_n \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
y \cdot x_1 \\
y \cdot x_2 \\
y \cdot x_n \\
\end{bmatrix}
\begin{bmatrix}
\xi_1 \\
\xi_2 \\
\xi_n \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
x_1 & 0 & \cdots & 0 \\
0 & x_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & x_n \\
\end{bmatrix}
\begin{bmatrix}
\xi_1 \\
\xi_2 \\
\xi_n \\
\end{bmatrix}
\]
where $yx$ is a column of scalar products between $y$ and the basis vectors \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\} and $X$ is a matrix of scalar products of the basis vectors.

If the basis vectors are orthogonal, $X = \text{diagonal} \left[ \frac{\langle y, x_i \rangle}{\langle x_i, x_i \rangle} \right]$

If the basis vectors are orthonormal, $X = \text{matrix} \left[ \frac{\langle y, x_i \rangle}{\langle x_i, x_i \rangle} \right]$

Combining the above $n$ projection operators in a row

$$\left\langle \mathbf{P}_1 \mathbf{y} \quad \mathbf{P}_2 \mathbf{y} \quad \ldots \quad \mathbf{P}_n \mathbf{y} \right\rangle = \left[ \begin{array}{c} \langle y, x_1 \rangle \\ \langle x_1, x_1 \rangle \\ \ldots \\ \langle x_n, x_n \rangle \end{array} \right] \begin{bmatrix} x_1 & 0 & \ldots & 0 \\ 0 & x_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & x_n \end{bmatrix}$$

$$\left\langle \mathbf{P}_1 \mathbf{y} \quad \mathbf{P}_2 \mathbf{y} \quad \ldots \quad \mathbf{P}_n \mathbf{y} \right\rangle = \left( \begin{array}{c} x_1 \\ 0 \end{array} \right)$$

$$\left\langle \mathbf{P}_1 \mathbf{y} \quad \mathbf{P}_2 \mathbf{y} \quad \ldots \quad \mathbf{P}_n \mathbf{y} \right\rangle = \left( \begin{array}{c} x_1 \\ 0 \end{array} \right)$$

where $\mathbf{y}^\mathsf{T}$ is a row of scalar products between $y$ and the basis vectors \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\} and $X$ is a matrix of scalar products of the basis vectors.
• Example Projection of a given vector $y$ in n-dimensional space onto another vector $x$. $x \in \{x_1, x_2, ..., x_n\}$.

$$\mathbb{P}_y \left( \frac{y \cdot x}{(x, x)} \right) x \quad \text{... projection of } y \text{ along } x$$

Any vector $y$ in n-dimensional LVS can be represented as a linear combination of the basis vectors

$$y = \xi_1 x_1 + \xi_2 x_2 + \ldots + \xi_n x_n \quad x = \psi_1 x_1 + \psi_2 x_2 + \ldots + \psi_n x_n$$

$$\mathbb{P}_y \left( \frac{y \cdot x}{(x, x)} \right) = \left[ \begin{array}{c}
\psi_1 x_1 + \psi_2 x_2 + \ldots + \psi_n x_n \\
\psi_1 x_1 + \psi_2 x_2 + \ldots + \psi_n x_n 
\end{array} \right] \left[ \begin{array}{c}
\xi_1 \\
\xi_2 \\
\ldots \\
\xi_n
\end{array} \right]$$

Addition|sum of two transformations $A$ & $B$

$$(A + B) \cdot x = A \cdot x + B \cdot x \quad C = A + B \quad \text{C matrix} = A \text{ matrix} + B \text{ matrix}$$

Multiplication of a transform $A$ by a scalar $\alpha$

$$\alpha \cdot A \cdot x = \alpha \cdot (A \cdot x) \quad C = \alpha \cdot A \quad \text{C matrix} = \alpha \cdot A \text{ matrix}$$

Multiple linear transformations in sequence:

$B$ followed by $A$  $A \cdot (B \cdot x) = (A \cdot B) \cdot x = C \cdot x \quad C = A \cdot B \quad \text{C matrix} = A \text{ matrix} \cdot B \text{ matrix}$

$A$ followed by $B$  $B \cdot (A \cdot x) = (B \cdot A) \cdot x = D \cdot x \quad D = B \cdot A \quad \text{D matrix} = B \text{ matrix} \cdot A \text{ matrix}$

In general, the order of operation matters:  $A \cdot B = B \cdot A \quad A \text{ matrix} \cdot B \text{ matrix} = B \text{ matrix} \cdot A \text{ matrix}$
Identity transformation. An identity transformation is a transformation that does not modify the given vector, i.e., simply do nothing!

\[ \mathbf{J} \cdot \mathbf{x} = \mathbf{x} \]

Thus, the action of the identity transformation on \( n \) linearly independent basis vectors \( \{x_1, x_2, \ldots, x_n\} \) is compactly expressed as: (Again, remember, some people abhor this compact notation.)

Rule for identity transformation:
\[
\begin{align*}
\mathbf{J} \cdot x_1 &= x_1 \\
\mathbf{J} \cdot x_2 &= x_2 \\
&\vdots \\
\mathbf{J} \cdot x_n &= x_n \\
\end{align*}
\]

compact matrix notation with vectors \( \mathbf{x} \) in a column
\[
\mathbf{J} \cdot \mathbf{x} = \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{I} \text{matrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}
\]

compact matrix notation with vectors \( \mathbf{x} \) in a row
\[
\mathbf{J} \cdot \mathbf{x} = \mathbf{y} = \begin{bmatrix} y_0 & y_1 & \cdots & y_n \end{bmatrix} = \begin{bmatrix} x_0 & x_1 & \cdots & x_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{I} \text{matrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix}
\]

Here, "\( \mathbf{J} \)" symbolizes identity transformation; whereas, "\( \mathbf{I} \text{matrix} \)" is an identity matrix that characterizes the identity transformation.

Inverse Transformation. Given a linear transformation \( \mathbf{A} \) in a LVS, an inverse transformation reverses the transformation performed on any vector \( \mathbf{x} \) by \( \mathbf{A} \). In other words, the linear transformation \( \mathbf{B} \) is an inverse of the transformation \( \mathbf{A} \) if, for every vector in the LVS, we have,

\[
\mathbf{A} \cdot \mathbf{B} \cdot \mathbf{x} = \mathbf{B} \cdot \mathbf{A} \cdot \mathbf{x} = \mathbf{x} \]

Note that both relationships have to hold, as \( \mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A} \) in general.

\[
\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A} = \mathbf{I}
\]

If \( \mathbf{B} \) is an inverse of \( \mathbf{A} \), then \( \mathbf{A} \) is an inverse of \( \mathbf{B} \). We commonly use the superscript "\( ^{-1} \)" to denote the inverse transformation of \( \mathbf{A} \): \( \mathbf{B} \) or \( \mathbf{A}^{-1} \). If inverse exists, it is unique. If a linear transformation \( \mathbf{A} \) has an inverse, it is said to be non-singular. Otherwise, it is singular. Again, this terminology is a very general one. We do not even have to be concerned about a real Euclidean space, finite dimension, matrix representation, etc.

If \( \mathbf{A} \) has an inverse, then \( \mathbf{A} \cdot \mathbf{x} = \mathbf{A} \cdot \mathbf{y} \) implies \( \mathbf{x} = \mathbf{y} \). In fact, this is a necessary condition for \( \mathbf{A} \) to be non-singular. In plain English, given two identical vectors \( \mathbf{z}_1 \) & \( \mathbf{z}_2 \) that result from the same linear transformation \( \mathbf{A} \), \( \mathbf{z}_1 = \mathbf{A} \cdot \mathbf{x} \) and \( \mathbf{z}_2 = \mathbf{A} \cdot \mathbf{y} \), if we trace back the linear transform by taking \( \mathbf{B} = \mathbf{A}^{-1} \) to see where each has originated from, we should end up at exactly the same starting vector (meaning \( \mathbf{x} = \mathbf{y} \)).
A common mis-conception. When we reverse the vector direction, we turn a vector \( \mathbf{x} \) into a reversed/inverted (i.e. negative) vector, and we denote this with a minus sign: \(-\mathbf{x}\). The process of reversing vector direction is not an inverse transform. Conversely, an inverse transformation is not \( \mathbf{A}^{-1} \). The notation \( \mathbf{A}^{-1} \) all by itself has no meaning; \(-\mathbf{A} \mathbf{x}\), which is a common abbreviation for \( -(\mathbf{A} \mathbf{x}) \) mathematically means the negative of the output vector \( \mathbf{A} \mathbf{x} \), just as \(-f(x)\), which is an abbreviation for \( -(f(x)) \), not \( (-f)(x) \), means the negative of the output from the function \( f \).

- **Example.** Rotations of arrows have inverses (i.e., the original rotation can be undone).
- **Example.** Column of numbers (e.g., a point/vector in Cartesian coordinate)

\[
\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}
\]

Do not confuse \( x_1 \) here (which is a number) with \( x_1 \) in the previous pages (which is a vector).

**Define** \( \mathbf{A} \mathbf{x} = \mathbf{y} \) in the usual way. \( \mathbf{y} = \mathbf{A} \mathbf{x} = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = \sum_{i=1}^{n} a_{ij}x_j \) for \( i = 1, 2, \ldots n \)

**Action of transformation** \( \mathbf{A} \) expressed in a compact matrix notation:

\[
\begin{bmatrix}
  y_1 \\
  y_2 \\
  \vdots \\
  y_n
\end{bmatrix} =
\begin{bmatrix}
  a_{11} & a_{12} & \ldots & a_{1n} \\
  a_{21} & a_{22} & \ldots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{n1} & a_{n2} & \ldots & a_{nn}
\end{bmatrix}
\begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_n
\end{bmatrix}
\]

The first ":" is linear operation; the second ":" is matrix multiplication.

We can think of \( x_1 \) etc as one of the basis vectors (rather than a number embedded in a column).

\[
\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}
\]

**compact matrix notation with vectors \( \mathbf{x} \) in a column (has the same form as the the linear algebraic equation but \( \mathbf{x}_1 \) & \( \mathbf{y}_1 \) in each column represents a basis vector).**

\[
\begin{bmatrix}
  y_1 \\
  y_2 \\
  \vdots \\
  y_n
\end{bmatrix} =
\begin{bmatrix}
  a_{11} & a_{12} & \ldots & a_{1n} \\
  a_{21} & a_{22} & \ldots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{n1} & a_{n2} & \ldots & a_{nn}
\end{bmatrix}
\begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_n
\end{bmatrix}
\]

**compact matrix notation with vectors \( \mathbf{x} \) in a row**

\[
\begin{bmatrix}
  y_1 \\
  y_2 \\
  \vdots \\
  y_n
\end{bmatrix} =
\begin{bmatrix}
  x_1 & x_2 & \ldots & x_n
\end{bmatrix}
\begin{bmatrix}
  a_{11} & a_{21} & \ldots & a_{n1} \\
  a_{12} & a_{22} & \ldots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{1n} & a_{2n} & \ldots & a_{nn}
\end{bmatrix}
\]

\[
\mathbf{x} \cdot \mathbf{A} \]
The linear transformation \( \mathcal{A} \) has inverse only if we can find \( x \) from given \( y \). Thus, \( \mathcal{A} \) is non-singular if and only if \( \det(\mathcal{A} \text{ matrix}) \neq 0 \).

**Example.** Row of numbers \( x = (x_1, x_2, \ldots, x_n) \):

Do not confuse "\( x_1 \)" here (which is a number) with "\( x \)" in the previous pages (which is a vector).

Define \( \mathcal{A}x = y \) in the non-usual way. \( y_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \), for \( i = 1, 2, \ldots n \).

In compact matrix notation:

\[
\begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} \\
  a_{21} & a_{22} & \cdots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{n1} & a_{n2} & \cdots & a_{nn}
\end{bmatrix}
\begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_n
\end{bmatrix}
= \begin{bmatrix}
  y_1 \\
  y_2 \\
  \vdots \\
  y_n
\end{bmatrix} = x \cdot \text{A matrix}
\]

**Example.** Projection of arrows has no inverse (i.e., singular). Reason: Many arrows yield the same projection; thus, we cannot get back the original arrow.

**Example.** All polynomials of degree \( n \) or less

Define two linear transformations \( \mathcal{A} \) and \( \mathcal{B} \):

\[
\mathcal{A}P = \frac{d}{dt}(P(t))
\]

\[
\mathcal{B}P = \frac{1}{t} \int_{0}^{t} P(\tau) \, d\tau
\]

Apply two linear transformations in sequence:

\[
\mathcal{A}(\mathcal{B}P) = \frac{d}{dt} \left( \frac{1}{t} \int_{0}^{t} P(\tau) \, d\tau \right) = \frac{d}{dt} \left( \frac{1}{t} \right) P(t) = P(t)
\]

\[
\mathcal{B}(\mathcal{A}P) = \frac{1}{t} \int_{0}^{t} \frac{d}{d\tau} (\mathcal{A}P(t)) \, d\tau = \frac{1}{t} \int_{0}^{t} \frac{d}{d\tau}(\mathcal{A}P(t)) \, d\tau = \frac{1}{t} \int_{0}^{t} \frac{d}{d\tau}(t \cdot P(t)) \, d\tau = P(t)
\]

Since \( \mathcal{A}(\mathcal{B}P) = \mathcal{B}(\mathcal{A}P) = P \), the above two linear transformations \( \mathcal{A} \) and \( \mathcal{B} \) are inverse of each other.

**Example:** Laplace Transform (Since the LVS is composed of continuous functions; the dimension is infinity, and there is no compact matrix representation of the action of the transformation on basis vectors.)

\[
\mathcal{A}f(t) = \int_{0}^{\infty} e^{-s \tau} \cdot f(t) \, d\tau = F(s)
\]

\[
\mathcal{A}^{-1}F(s) = \frac{1}{2\pi i} \int_{y-i\infty}^{y+i\infty} e^{s\tau} \cdot F(s) \, ds = f(t)
\]
Example: Fourier Sine Transform.
\[
\mathcal{F}\cdot f(\omega) = \int_{\omega}^{\infty} \sin(\omega \cdot t) \cdot f(t) \, dt = F(\omega)
\]
\[
F^{-1}(\omega) = \int_{0}^{\omega} \sin(\omega \cdot t) \cdot F(\omega) \, dt = f(t)
\]

Example: Fourier Cosine Transform.
\[
\mathcal{F}\cdot f(\omega) = \int_{\omega}^{\infty} \cos(\omega \cdot t) \cdot f(t) \, dt = F(\omega)
\]
\[
F^{-1}(\omega) = \int_{0}^{\omega} \cos(\omega \cdot t) \cdot F(\omega) \, dt = f(t)
\]

Example: Fourier Transform.
\[
\mathcal{F}\cdot f(t) = \int_{-\infty}^{\infty} e^{i \cdot 2 \pi \cdot \omega \cdot t} \cdot f(t) \, dt = F(\omega)
\]
\[
F^{-1}(\omega) = \int_{-\infty}^{\infty} e^{-i \cdot 2 \pi \cdot \omega \cdot t} \cdot F(\omega) \, dt = f(t)
\]

**Null Space.** The null space of a linear transformation \( \mathcal{A} \) is the set of all vectors \( x \) in the LVS such that
\[
\mathcal{A} \cdot x = 0
\]

Naturally, these special set of vectors \( x \) lie within the same LVS in which the linear transformation \( \mathcal{A} \) is defined. Thus, the null space is a subspace of the LVS. Note that \( \mathcal{A} \) does not need to have an inverse in order for us to speak of the null space.
Orthogonal Transformation. An orthogonal transformation is a linear transformation $\mathcal{A}$ that acts on a vector $x$ and does not change its length. In other words, it is a transformation that merely "rotate" a vector (thus, change its angle but not its length). A common mis-conception: an orthogonal transformation does not mean the input vector $x$ and the output vector $y=\mathcal{A}x$ are orthogonal to each other! In other words, the angle of "rotation" between the input vector $x$ and the output vector $y=\mathcal{A}x$ does not need to be 90 degrees. Another common mis-conception: do not confuse orthogonal vectors with orthogonal linear transformations! Basis vectors that we use to describe a given vector in LVS do not need to be orthogonal. Thus, we can talk about orthogonal transformations in LVS that is represented by non-orthogonal basis vectors. Conversely, in a LVS represented by orthogonal basis vectors, the linear transformation can be anything (certainly not limited by orthogonal transformations). Since we are referring to vector length, we need to define a scalar product and we are working with real Euclidean space.

Definition of orthogonal transformation $|\mathcal{A}x| = |x|

Scalar product after transformation = scalar product before transformation

$(\mathcal{A}x, \mathcal{A}y) = (x, y)$

"Orthogonal" is a way to describe the "before" and the "after" of a linear transformation $\mathcal{A}$ on $x$. An orthogonal transformation preserves scalar product (not just for $x$, but for any $x$ & $y$ pairs).

$(\mathcal{A}x, \mathcal{A}y) = (x, y)$

Proof. Let $z$ be the sum of two vectors $x$ & $y$. $z = x + y$

Apply transformation $\mathcal{A}$ to $z$; since $\mathcal{A}$ is orthogonal, by definition of $\mathcal{A}$ being "orthogonal" we have for $z$

$(\mathcal{A}z, \mathcal{A}z) = (z, z)$

We substitute $z$ into the above definition.

$(\mathcal{A}(x + y), \mathcal{A}(x + y)) = ((x + y), (x + y))$

$(\mathcal{A}x, \mathcal{A}y) + (\mathcal{A}x, \mathcal{A}y) + (\mathcal{A}y, \mathcal{A}x) + (\mathcal{A}y, \mathcal{A}y) = (x, x) + (x, y) + (y, x) + (y, y)$

Again, by definition of $\mathcal{A}$ being "orthogonal" we have for $x$ & $y$: $(\mathcal{A}x, \mathcal{A}x) = (x, x)$ $(\mathcal{A}y, \mathcal{A}y) = (y, y)$

$(x, x) + 2(\mathcal{A}x, \mathcal{A}y) + (y, y) = (x, x) + 2(x, y) + (y, y)$

Comparing the LHS & RHS, we have

$(\mathcal{A}x, \mathcal{A}y) = (x, y)$

Action of $\mathcal{A}$ on each of the basis vectors $\{x_1, x_2, \ldots, x_n\}$

- Action of $\mathcal{A}$ on $x_1$
  
  $\mathcal{A}x_1 = a_1x_1 + a_2x_2 + \ldots + a_nx_n$

- Action of $\mathcal{A}$ on $x_2$
  
  $\mathcal{A}x_2 = a_1x_2 + a_2x_2 + \ldots + a_nx_n$

- ... 

- Action of $\mathcal{A}$ on $x_n$
  
  $\mathcal{A}x_n = a_1x_n + a_2x_n + \ldots + a_nx_n$

Action of transformation $\mathcal{A}$ expressed in a compact matrix notation with vectors $x$ & $y$ in a column

$$\mathcal{A}x = \begin{bmatrix} \mathcal{A}x_1 \\ \mathcal{A}x_2 \\ \vdots \\ \mathcal{A}x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \mathcal{A}_{\text{matrix}} x$$

$$y = \mathcal{A}_{\text{matrix}} x$$
compact matrix notation with vectors \( x \) & \( y \) in a row

\[
\mathcal{A}x = \begin{pmatrix} \mathcal{A}x_1 \\ \mathcal{A}x_2 \\ \vdots \\ \mathcal{A}x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}
\]

Matrix of various scalar products between basis vectors \((\mathcal{A}x_i, \mathcal{A}x_j)\) \(= (x_i, x_j)\).

\[
\begin{pmatrix}
(\mathcal{A}x_1, \mathcal{A}x_1) & (\mathcal{A}x_1, \mathcal{A}x_2) & \cdots & (\mathcal{A}x_1, \mathcal{A}x_n) \\
(\mathcal{A}x_2, \mathcal{A}x_1) & (\mathcal{A}x_2, \mathcal{A}x_2) & \cdots & (\mathcal{A}x_2, \mathcal{A}x_n) \\
\vdots & \vdots & \ddots & \vdots \\
(\mathcal{A}x_n, \mathcal{A}x_1) & (\mathcal{A}x_n, \mathcal{A}x_2) & \cdots & (\mathcal{A}x_n, \mathcal{A}x_n)
\end{pmatrix}
= \begin{pmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{pmatrix}
\]

\(A_{\text{matrix}}^T \cdot X \cdot A_{\text{matrix}} = X \quad \text{where} \quad X \text{ is a matrix of scalar products of the basis vectors}\)

If \(\mathcal{A}\) is an orthogonal linear transformation, by definition of "orthogonal", the above relationship holds; it follows the similarity transform equation between \(A_{\text{matrix}}\) and the inverse of its transpose, \((A_{\text{matrix}}^T)^{-1}\), with the matrix of scalar products of basis \(X\) as the similarity transform matrix.

\[
A_{\text{matrix}}^T \cdot X \cdot A_{\text{matrix}} = \begin{pmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{pmatrix}
\]

\(\begin{pmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{pmatrix}
\)

\[
\begin{pmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{pmatrix}
= \begin{pmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{pmatrix}
\]

\(A_{\text{matrix}}^T \cdot X \cdot A_{\text{matrix}} = \begin{pmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{pmatrix}
\]

\(A_{\text{matrix}}^T \cdot \begin{pmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{pmatrix} = X = \begin{pmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{pmatrix}
\)

### Example
**Rotation** in a plane with \(\phi\)

\[
x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}
\]

Define a linear transformation \(\mathcal{A}\)

\[
\mathcal{A}x = \begin{pmatrix} \cos(\phi) \cdot x_1 - \sin(\phi) \cdot x_2 \\ \sin(\phi) \cdot x_1 + \cos(\phi) \cdot x_2 \end{pmatrix}
\]

Define scalar product \((x, y) = x_1 \cdot y_1 + x_2 \cdot y_2\)

\[
(x) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}
\]

\[
(\mathcal{A}x) = \begin{pmatrix} \cos(\phi) \cdot x_1 - \sin(\phi) \cdot x_2 \\ \sin(\phi) \cdot x_1 + \cos(\phi) \cdot x_2 \end{pmatrix}
\]

\[
(x) \cdot (\mathcal{A}x) = (x_1)^2 + (x_2)^2
\]

Thus, \(\mathcal{A}\) is an orthogonal transformation.

Show that the scalar product is preserved.
\[
\begin{align*}
(A \cdot x, A \cdot y) &= \left(\cos(\phi) \cdot x_1 - \sin(\phi) \cdot x_2, \cos(\phi) \cdot y_1 - \sin(\phi) \cdot y_2\right) \\
&\quad + \left(\sin(\phi) \cdot x_1 + \cos(\phi) \cdot x_2, \sin(\phi) \cdot y_1 + \cos(\phi) \cdot y_2\right) \\
&= x_2 \cdot y_2 + x_1 \cdot y_1 \equiv (x, y)
\end{align*}
\]

**Symmetric Transformation.** A symmetric transformation is a linear transformation \(A\) with the following property for every pair of vectors \(x\) and \(y\) in real Euclidean space.

 definition of symmetric transformation \((A \cdot x, y) \equiv (x, A \cdot y)\)

In other words, given any two vectors \(x\) and \(y\), swapping the operation, whether \(A\) acts on \(x\) or on \(y\), does not change the resulting scalar product between them, meaning the relationship (e.g., angle). A common mis-conception: a symmetric transformation does not mean the input vector \(x\) is symmetric, nor the output from the transformation \(y=A \cdot x\) is symmetric! Another common mis-conception: do not confuse orthogonal basis vectors with symmetric linear transformations! Basis vectors that we use to describe a given vector in LVS do not need to be orthogonal; they can be anything as long as they are linearly independent. Thus, we can talk about symmetric transformations in LVS that is represented by non-orthogonal basis vectors. Since we are referring to scalar product, the same transformation can be symmetric with respect to one scalar product definition, but not symmetric with a different scalar product definition. This is analogous to the orthogonality between two vectors: two same vectors are orthogonal with one scalar product definition, but non-orthogonal with a different scalar product definition.

Basically, it is an "equal opportunity" transform. "Symmetrical" is a way to describe that applying linear transformation \(A\) to a vector \(x\), in terms of its relationship to all the other vectors \(y\) as quantified by the scalar product \((A \cdot x, y)\), does not depend on which vector \(x\) \(A\) acts on. Had \(A\) acted on a different vector \(y\), the resulting scalar product \((A \cdot y, x)\) would have been identical to it acting on \(x\), i.e., \((A \cdot x, y)\). The change brought about by \(A\) acting on a random vector \(x\) is independent of the vector \(x\) that we put through \(A\). Each and every vector \(x\) is capable of causing the same result. "Orthogonal" deals with "before" and "after"; whereas, "symmetrical" deals with laterally swapping an input vector \(x\) with another input vector \(y\).

**Action of \(A\) on each of the basis vectors \(\{x_1, x_2, ..., x_n\}\)**

- **action of \(A\) on \(x_1\)**
  \[A \cdot x_1 = y_1 = a_{11} \cdot x_1 + a_{12} \cdot x_2 + ... + a_{1n} \cdot x_n\]

- **action of \(A\) on \(x_2\)**
  \[A \cdot x_2 = y_2 = a_{21} \cdot x_1 + a_{22} \cdot x_2 + ... + a_{2n} \cdot x_n\]

- **action of \(A\) on \(x_n\)**
  \[A \cdot x_n = y_n = a_{n1} \cdot x_1 + a_{n2} \cdot x_2 + ... + a_{nn} \cdot x_n\]

**Action of transformation \(A\) expressed in a compact matrix notation with vectors \(x\) & \(y\) in columns**

\[
\begin{bmatrix}
A \cdot x_1 \\
A \cdot x_2 \\
\vdots \\
A \cdot x_n
\end{bmatrix} = 
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n
\end{bmatrix} = 
\begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix} = A \text{ matrix} \cdot x \\
\begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{bmatrix}
\]

\(y = A \text{ matrix} \cdot x\)
compact matrix notation with vectors \( x \) & \( y \) in rows

\[
A \cdot x = \begin{bmatrix} A \cdot x_1 & A \cdot x_2 & \cdots & A \cdot x_n \end{bmatrix} \cdot \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \cdot \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{11} & a_{22} & \cdots & a_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix}
\]

Matrix of various scalar products between basis vectors and the output of \( A \) on basis vectors \((A x_1, x_2) = (x_1, A x_2)\).

\[
\begin{bmatrix} A \cdot x_1, x_1 & A \cdot x_1, x_2 & \cdots & A \cdot x_1, x_n \\ A \cdot x_2, x_1 & A \cdot x_2, x_2 & \cdots & A \cdot x_2, x_n \\ \vdots & \vdots & \ddots & \vdots \\ A \cdot x_n, x_1 & A \cdot x_n, x_2 & \cdots & A \cdot x_n, x_n \end{bmatrix}
= A_{\text{matrix}} \cdot X
\]

\[
\begin{bmatrix} x_1, A \cdot x_1 & x_1, A \cdot x_2 & \cdots & x_1, A \cdot x_n \\ x_2, A \cdot x_1 & x_2, A \cdot x_2 & \cdots & x_2, A \cdot x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_n, A \cdot x_1 & x_n, A \cdot x_2 & \cdots & x_n, A \cdot x_n \end{bmatrix}
= X \cdot A_{\text{matrix}}^T
\]

\(A_{\text{matrix}}\) is symmetric, by definition of "symmetric", the above two equations are equal. The resulting relationship follows the similarity transform equation between \( A_{\text{matrix}} \) and its transpose \( A_{\text{matrix}}^T\), with the matrix of scalar products of basis \( X \) as the similarity transform matrix.

\[
A_{\text{matrix}} \cdot X = X \cdot A_{\text{matrix}}^T
\]

for general basis vectors that are not necessarily orthogonal (basis need to be linearly independent, but do not need to be orthogonal)

\[
A_{\text{matrix}} = A_{\text{matrix}}^T
\]

for orthonormal basis vectors \( \{x_1, x_2, \ldots, x_n\} \)

\( X = I_{\text{matrix}} \)

**Example.** columns of numbers

\[
\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}
\]

Define a linear transformation \( A \)

\[
A \cdot x = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}
\]

Define scalar product \( (x, y) = x_1 y_1 + x_2 y_2 \)

\[
0 = (A \cdot x, y) - (x, A \cdot y)
\]

requirement for a symmetric transformation.

\[
0 = \left( a_{11} x_1 + a_{12} x_2, y_1 + a_{21} x_1 + a_{22} x_2 \right) - \left( x_1 \left( a_{11} y_1 + a_{12} y_2 \right) + x_2 \left( a_{21} y_1 + a_{22} y_2 \right) \right)
\]

\[
0 = \left( a_{12} - a_{21} \right) \cdot (y_1 x_2 - x_1 y_2)
\]

Thus, the requirement for a symmetrical transformation is \( a_{12} = a_{21} \)
**Example. Convolution Integral.** Let the LVS be all continuous functions in $t= [0, 1]$

Define a linear transformation $A$ (convolution integral), where $K$ is a continuous function $K(t, \tau)$. 

$$A: f = \int_0^1 K(t, \tau) \cdot f(\tau) \, d\tau$$

Define scalar product $(f,g) = \int_0^1 f(\tau) \cdot g(\tau) \, d\tau$

$$0 = (Af, g) - (f, Ag) \quad \text{... requirement for a symmetric transformation.}$$

$$0 = \int_0^1 \left( \int_0^1 K(t, \tau) \cdot f(\tau) \, d\tau \right) g(t) \, dt - \int_0^1 f(t) \cdot \left( \int_0^1 K(t, \tau) \cdot g(\tau) \, d\tau \right) \, dt$$

$$0 = \int_0^1 \int_0^1 K(t, \tau) \cdot f(\tau) \cdot g(t) \, d\tau \, dt - \int_0^1 \int_0^1 K(t, \tau) \cdot f(\tau) \cdot g(t) \, d\tau \, dt$$

$$0 = \int_0^1 \int_0^1 (K(t, \tau) - K(\tau, t)) \cdot f(\tau) \cdot g(t) \, d\tau \, dt$$

Thus, the requirement for a symmetrical transformation is $K(t, \tau) = K(\tau, t)$

**Skew-Symmetric Transformation.** A skew-symmetric transformation is a linear transformation $A$ with the following property for every pair of vectors $x$ and $y$ in real Euclidean space.

$$(Ax, y) = -(x, Ay)$$

compact matrix notation

$$A_{\text{matrix}}^T \cdot X = -X \cdot A_{\text{matrix}}^T \quad \text{... for general basis vectors not necessarily orthogonal}$$

$$A_{\text{matrix}}^T = -A_{\text{matrix}}^T \quad \text{... for orthonormal basis vectors} \{x_1, x_2, ..., x_n\} \quad a_{ij} = -a_{ji}$$
Transpose. Any linear transformation \( A \) in n-dimensional LVS has a unique transpose. The transpose of a linear transformation \( A \) has the following property for every pair of vectors \( x \) and \( y \) in real Euclidean space. \( A^T \) is the transpose of \( A \), and \( A \) is the transpose of \( A^T \). ("the", and not "a", means that the transpose of \( A \) exists and it is unique.) Notation: \( A^T \).

\[
(A \cdot x, y) = (x, A^T \cdot y)
\]

Special case. If \( A \) is a symmetrical linear transformation, by definition we have, \((A \cdot x, y) = (x, A \cdot y)\) \(A^T = A \) ... for symmetrical linear transformation

Some properties of transpose. (including, but not limited to, the traditional "matrix" type)

\[
(A + B)^T = A^T + B^T
\]

\[
(\alpha A)^T = \alpha A^T
\]

\[
(A \cdot B)^T = B^T \cdot A^T
\]

Action of \( A \) on each of the basis vectors \( \{x_1, x_2, ..., x_n\} \)

Action of \( A \) on \( x_1 \)
\[
A \cdot x_1 = y_1 = a_{11} \cdot x_1 + a_{21} \cdot x_2 + ... + a_{1n} \cdot x_n
\]

Action of \( A \) on \( x_2 \)
\[
A \cdot x_2 = y_2 = a_{21} \cdot x_1 + a_{22} \cdot x_2 + ... + a_{2n} \cdot x_n
\]

Action of \( A \) on \( x_n \)
\[
A \cdot x_n = y_n = a_{n1} \cdot x_1 + a_{n2} \cdot x_2 + ... + a_{nn} \cdot x_n
\]

Action of transformation \( A \) expressed in a compact matrix notation with vectors \( x \) & \( y \) in a column

\[
\begin{bmatrix}
A \cdot x_1 \\
A \cdot x_2 \\
... \\
A \cdot x_n
\end{bmatrix} =
\begin{bmatrix}
y_1 \\
y_2 \\
... \\
y_n
\end{bmatrix} =
\begin{bmatrix}
a_{11} & a_{12} & ... & a_{1n} \\
a_{21} & a_{22} & ... & a_{2n} \\
... & ... & ... & ... \\
a_{n1} & a_{n2} & ... & a_{nn}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
... \\
x_n
\end{bmatrix}
\]

\[
A = A_{\text{matrix}} \cdot x
\]

Action of \( A \) expressed in a compact matrix notation with vectors \( x \) & \( y \) in a row

\[
\begin{bmatrix}
A \cdot x_1 & A \cdot x_2 & ... & A \cdot x_n
\end{bmatrix} =
\begin{bmatrix}
y_1 & y_2 & ... & y_n
\end{bmatrix} =
\begin{bmatrix}
a_{11} & a_{12} & ... & a_{1n} \\
a_{21} & a_{22} & ... & a_{2n} \\
... & ... & ... & ... \\
a_{n1} & a_{n2} & ... & a_{nn}
\end{bmatrix}
\begin{bmatrix}
x_1 & x_2 & ... & x_n
\end{bmatrix}
\]

\[
A = A_{\text{matrix}} \cdot x
\]

Matrix of various scalar products between basis vectors \( \langle A \cdot x_i, x_j \rangle = \langle x_i, A^T \cdot x_j \rangle \).

\[
\begin{bmatrix}
\langle A \cdot x_1, x_1 \rangle & \langle A \cdot x_1, x_2 \rangle & ... & \langle A \cdot x_1, x_n \rangle \\
\langle A \cdot x_2, x_1 \rangle & \langle A \cdot x_2, x_2 \rangle & ... & \langle A \cdot x_2, x_n \rangle \\
... & ... & ... & ... \\
\langle A \cdot x_n, x_1 \rangle & \langle A \cdot x_n, x_2 \rangle & ... & \langle A \cdot x_n, x_n \rangle
\end{bmatrix} =
\begin{bmatrix}
a_{11} & a_{12} & ... & a_{1n} \\
a_{21} & a_{22} & ... & a_{2n} \\
... & ... & ... & ... \\
a_{n1} & a_{n2} & ... & a_{nn}
\end{bmatrix}
\begin{bmatrix}
x_1 & x_2 & ... & x_n
\end{bmatrix}
\]

\[
= A_{\text{matrix}} \cdot x
\]
\[
\left[
\begin{array}{ccc}
\langle x_1, A^T x_1 \rangle & \langle x_1, A^T x_2 \rangle & \cdots & \langle x_1, A^T x_n \rangle \\
\langle x_2, A^T x_1 \rangle & \langle x_2, A^T x_2 \rangle & \cdots & \langle x_2, A^T x_n \rangle \\
& \cdots & \cdots & \cdots \\
\langle x_n, A^T x_1 \rangle & \langle x_n, A^T x_2 \rangle & \cdots & \langle x_n, A^T x_n \rangle \\
\end{array}
\right] =
\left[
\begin{array}{ccc}
\langle x_1, x_1 \rangle & \langle x_1, x_2 \rangle & \cdots & \langle x_1, x_n \rangle \\
\langle x_2, x_1 \rangle & \langle x_2, x_2 \rangle & \cdots & \langle x_2, x_n \rangle \\
& \cdots & \cdots & \cdots \\
\langle x_n, x_1 \rangle & \langle x_n, x_2 \rangle & \cdots & \langle x_n, x_n \rangle \\
\end{array}
\right]
\left[
\begin{array}{cccc}
a'_{11} & a'_{12} & \cdots & a'_{1n} \\
a'_{12} & a'_{22} & \cdots & a'_{2n} \\
& \cdots & \cdots & \cdots \\
a'_{1n} & a'_{2n} & \cdots & a'_{nn} \\
\end{array}
\right]
\]

where \( X \) is a matrix of scalar products of the basis vectors and

\[ B = X A' \] describes the action of transpose of \( A \) on each of the basis vectors \( \{x_1, x_2, \ldots, x_n\} \)

\[ A'A = X' \] for general basis vectors that are not necessarily orthogonal (basis need to be linearly independent, but do not need to be orthogonal)

\[ A = X'B \] for orthonormal basis vectors \( \{x_1, x_2, \ldots, x_n\} \)

Because we can always find an orthonormal set of basis vectors in \( n \)-dimensional space (Gram-Schmidt assures us that this can always be done), a linear transform \( A \) can act on the basis \( \{x_1, x_2, \ldots, x_n\} \), then we can always define another linear transform \( B = A^T \) that acts on the basis \( \{x_1, x_2, \ldots, x_n\} \) according to the above rule. Thus, irrespective of the linear transformation (no need to be orthogonal or symmetric, etc), there is always a unique transpose \( B = A^T \).