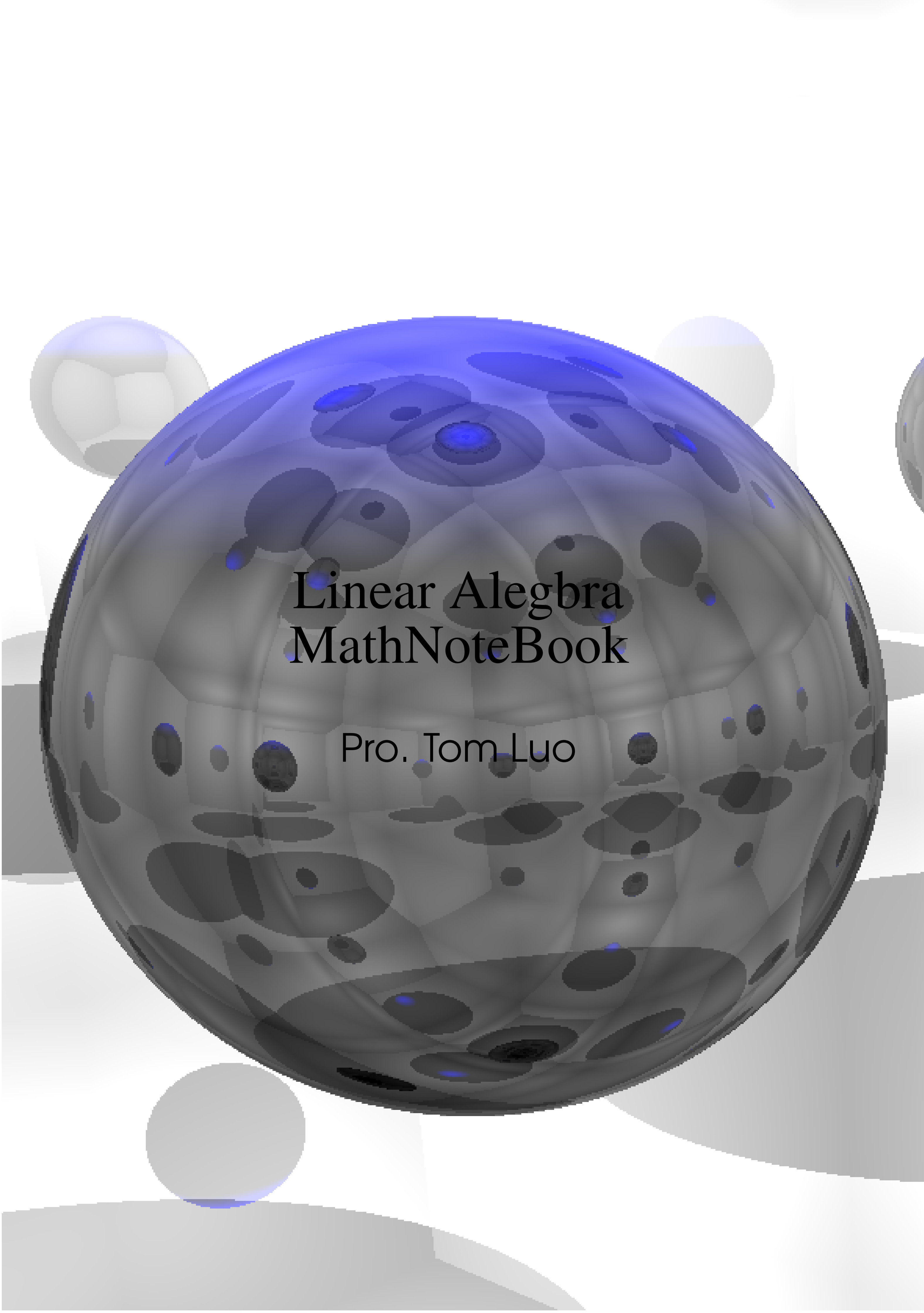


Linear Alegbra MathNoteBook

Pro. Tom Luo

10 — Week4

10.1 Friday

10.1.1 Linear Transformation

Start with a matrix $\boldsymbol{A}$. When multiplying $\boldsymbol{A}$ with a vector $\boldsymbol{v}$, it transform $\boldsymbol{v}$ to another vector $\boldsymbol{A}\boldsymbol{v}$. Matrix multiplication $L(\boldsymbol{v}) = \boldsymbol{A}\boldsymbol{v}$ gives a **linear transformation**:

Definition 10.1 — linear transformation. A transformation L assigns an output $T(\boldsymbol{v})$ to each inpout vector $\boldsymbol{v}$ in $\boldsymbol{V}$.
The transformation is **linear transformation** if it satisfies

$$L(\alpha \boldsymbol{v}_1 + \beta \boldsymbol{v}_2) = \alpha L(\boldsymbol{v}_1) + \beta L(\boldsymbol{v}_2)$$

for all vector v_1, v_2 and scalar α, β.

Key Observation: If the input is $\boldsymbol{v} = \boldsymbol{0}$, the output must be $L(\boldsymbol{v}) = \boldsymbol{0}$.

The idea of linear transformation

Given linear transformation $L : \mathbb{R}^n \mapsto \mathbb{R}^m$, let's show that in order to study the output we only need to start from the **basis** of our output:
Assume the basis of $\mathbb{R}^n$ is $\{e_1, e_2, \dots, e_n\}$, where $L(e_i) = a_i \in \mathbb{R}^m$ for $i = 1, \dots, n$.
Notice that **The rule of linearity extends to combinations of three vectors or $\boldsymbol{n}$ vectors.**
Hence given any vector $\boldsymbol{x} = x_1e_1 + x_2e_2 + \cdots + x_ne_n \in \mathbb{R}^n$, we express its transformation in matrix multiplication form:

$$\begin{aligned}
L(\boldsymbol{x}) &= L(x_1e_1 + x_2e_2 + \cdots + x_ne_n) \\
&= x_1L(e_1) + x_2L(e_2) + \cdots + x_nL(e_n) \\
&= x_1a_1 + x_2a_2 + \cdots + x_na_n = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} \\
&= \boldsymbol{A}\boldsymbol{x}
\end{aligned}$$

where $\boldsymbol{A}$ is a $m \times n$ matrix with columns $a_1, \dots, a_n$.

Matrix defines linear transformation

Conversely, given $m\times n$ matrix $\boldsymbol{A}$, $L(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$ defines a linear mapping. This is because matrix multiplication is also a lienar operator.

R Transformations have a new "language". For example, for nonlinear transformation, if there is **no matrix**, we cannot talk about a **column space**. But this idea could be rescued. We know the *column space* consists of all outputs $\boldsymbol{A}\boldsymbol{v}$, the *nullspace* consists of all inputs for which $\boldsymbol{A}\boldsymbol{v} = \boldsymbol{0}$. We could translate those terms into "range" and "kernel":

Definition 10.2 — range. For a linear transformation $L: V \mapsto W$, the range (or image) of L refers to the set of all outputs $T(\boldsymbol{v})$, which is denoted as:

$$\text{Range}(L) = \{L(\boldsymbol{x}) : x \in \boldsymbol{V}\}$$

Sometimes we also use notation $\text{Im}(L)$ to express the same thing. ■

The range corresponds to column space. If $L(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$, we have $\text{Range}(L) = \text{col}(\boldsymbol{A})$.

Definition 10.3 — kernel. The kernel of L refers to the set of all inputs for which $L(\boldsymbol{v}) = \boldsymbol{0}$, which is denoted as:

$$\ker(L) = \{\boldsymbol{x} : L(\boldsymbol{x}) = \boldsymbol{0}\}$$

■

Kernel corresponds to nullspace. If $L(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$, we have $\ker(L) = N(\boldsymbol{A})$.

R For linear transformation $L: \boldsymbol{V} \mapsto \boldsymbol{W}$, where $L(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$. We have two rules:

$$\begin{aligned} N(\boldsymbol{A}) &\mapsto \{\boldsymbol{0}\} \\ \boldsymbol{V} &\mapsto \text{col}(\boldsymbol{A}) \end{aligned}$$

10.1.2 Example: differentiation

Key idea of this section:

Suppose we know $L(\boldsymbol{v}_1),\dots,L(\boldsymbol{v}_n)$ for the basis vectors $\boldsymbol{v}_1,\dots,\boldsymbol{v}_n$,Then the linearity produces $L(\boldsymbol{v})$ for every other input vector $\boldsymbol{v}$.

Reason: Every $\boldsymbol{v}$ is a unique combination $c_1\boldsymbol{v}_1+\cdots+c_n\boldsymbol{v}_n$ of the basis vector $\boldsymbol{v}_i$. Suppose L is a linear transformation, $L(\boldsymbol{v})$ must be the **same combination** $c_1L(\boldsymbol{v}_1)+\cdots+c_nL(\boldsymbol{v}_n)$ of the **known outputs** $L(\boldsymbol{v}_i)$.

The derivative of the functions $1,x,x^2,x^3$ are $0,1,2x,3x^2$. If we consider "**taking the derivative**" as a transformation, whose inputs and outputs are functions, then we claim that the **derivative transformation** is **linear**:

$$L(\boldsymbol{v}) = \frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}x} \qquad \text{obeys the linearity rule} \qquad \frac{\mathrm{d}}{\mathrm{d}x}(c\boldsymbol{v}+d\boldsymbol{w}) = c\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}x}+d\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}x}$$

If we consider $1,x,x^2,x^3$ as vectors instead of functions, we notice they form a basis for the space $\boldsymbol{V}$ of *polynomials with degree*≤ 3. Find derivatives of these four basis tells us all derivatives in $\boldsymbol{V}$:

■ **Example 10.1** Given any vector $\boldsymbol{v}$ in $\boldsymbol{V}$, it can be expressed as $\boldsymbol{v} = a + bx + cx^2 + dx^3$. Thus we want to find the derivative transformation output for $\boldsymbol{v}$:

$$\begin{aligned} L(\boldsymbol{v}) &= aL(1) + bL(x) + cL(x^2) + dL(x^3) \\ &= a \times (0) + b \times (1) + c \times (2x) + d \times (3x^2) \\ &= b + 2cx + 3dx^2 \end{aligned}$$

Can we express this linear transformation L by a matrix $\boldsymbol{A}$? The answer is *Yes*:
The derivative transforms the space $\boldsymbol{V}$ of cubics to the space $\boldsymbol{W}$ of quadratics. The basis for $\boldsymbol{V}$ is $1, x, x^2, x^3$. The basis for $\boldsymbol{W}$ is $1, x, x^2$. *The derivative matrix is 3 by 4*:

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} = \textbf{matrix form of derivative } L.$$

Why is $\boldsymbol{A}$ the correct matrix? Because **multiplying by $\boldsymbol{A}$ agrees with transforming by** L. The derivative of $\boldsymbol{v} = a + bx + cx^2 + dx^3$ is $L(\boldsymbol{v}) = b + 2cx + 3dx^2$. The same numbers b and $2c$ and $3d$ appear when we multiply by matrix $\boldsymbol{A}$:

$$\textbf{\textit{Take the derivative}} \qquad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} b \\ 2c \\ 3d \end{bmatrix}.$$

What does the matrix $\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$ and $\begin{bmatrix} b \\ 2c \\ 3d \end{bmatrix}$ mean?
It is the **coordinate vector** of $\boldsymbol{v}$ and $L(\boldsymbol{v})$. If we consider $a + bx + cx^2 + dx^3$ as a vector, then it's better for us to study its coordinate vector $\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$.
Hence taking derivative of $\boldsymbol{v}$ is the same as multiplying matrix $\boldsymbol{A}$ by its coordinate vector. ■

The inverse of the derivative.
The **integral** is the inverse of the derivative. That is the Fundamental Theorem of Calculus. We see it now in linear Algebra. The integral transformation L^{-1} that *takes the integral from 0 to x* is linear! Applying L^{-1} to $1, x, x^2$, which are $\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3$:

$$\textbf{Integration is } L^{-1} \qquad \int_0^x 1\,\mathrm{d}x = x, \quad \int_0^x x\,\mathrm{d}x = \frac{1}{2}x^2, \quad \int_0^x x^2\,\mathrm{d}x = \frac{1}{3}x^3.$$

By linearity, the integral of $\boldsymbol{w} = B + Cx + Dx^2$ is $L^{-1}(\boldsymbol{w}) = Bx + \frac{1}{2}Cx^2 + \frac{1}{3}Dx^3$. The integral of a quadratic is a cubic. The input space of L^{-1} is the quadratics, the output space is the cubics. **Integration takes W back to V**. Integration matrix will be 4 by 3:

$$\textbf{Take the integral} \qquad \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} B \\ C \\ D \end{bmatrix} = \begin{bmatrix} 0 \\ B \\ \frac{1}{2}C \\ \frac{1}{3}D \end{bmatrix}.$$

If our input is $\boldsymbol{w} = B + Cx + Dx^2$, our integral is $0 + Bx + \frac{1}{2}Cx^2 + \frac{1}{3}Dx^3$.

Recall we have express derivative and integral as matrix:

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \qquad \boldsymbol{A}^{-1} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$$

I want to call this matrix $\boldsymbol{A}^{-1}$, though rectangular matrices don't have inverses. We notice that $\boldsymbol{A}^{-1}$ is the **right inverse** of matrix $\boldsymbol{A}$! (Do you remember the definition that shown in mid-term?)

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{but} \qquad \boldsymbol{A}^{-1}\boldsymbol{A} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

This is reasonable. If you integrate a function and then differentiate, you get back to the start. Hence $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}$. But if you differentiate before integrating, the constant term is lost.
The integral of the derivative of 1 is zero:

$$L^{-1}L(1) = \text{integral of zero function} = 0.$$

Summary: In this example, we want to take the derivative. Then we let $\boldsymbol{V}$ be a vector space of polynomials with degree ≤ 3. Then its basis is given by $E = \{1, x, x^2, x^3\}$. Any $v \in \boldsymbol{V}$ there is a unique linear combination of the basis vectors that equals to v:

$$v = a + bx + cx^2 + dx^3$$

We write the coordinate vector of v relative to E:

$$[v]_E = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

Then we postmultiply $\boldsymbol{A}$ by $[v]_E$ to get the coordinate vector of output space:

$$[L(v)]_F = \boldsymbol{A}[v]_E$$

where $F = \{1, x, x^2\}$.
Here we give the formal definition for coordinate vector:

Definition 10.4 — coordinate vector. Let $\boldsymbol{V}$ be a vector space of dimension n and let $B = \{v_1, v_2, \dots, v_n\}$ be an **ordered** basis for $\boldsymbol{V}$. Then for any $v \in \boldsymbol{V}$ there is a unique linear combination of the basis vectors such that

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$$

where $\alpha_1, \dots, \alpha_n$ are scalars.
The **coordinate vector** of v relative to B is given by

$$[v]_B = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

Hence vector v could be expressed as: $v = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \times [v]_B$. ■

Also, there follows a theorem which is easy to verify:

Theorem 10.1 Let $E = \{v_1, \dots, v_n\}$ be a basis for $\boldsymbol{V}$; $F = \{w_1, \dots, w_m\}$ be a basis for $\boldsymbol{W}$. Given linear transformation $L : \boldsymbol{V} \mapsto \boldsymbol{W}$, for any vector $v \in \boldsymbol{V}$, there exists $m \times n$ matrix $\boldsymbol{A}$ such that

$$[L(v)]_F = \boldsymbol{A}[v]_E$$

And there is a corollary that is more commonly useful:

Corollary 10.1 Given linear transformation $L : \boldsymbol{V} \mapsto \boldsymbol{V}$. We set $E = \{\alpha_1, \dots, \alpha_n\}$ to be its basis. Then given any vector v, there exists $n \times n$ matrix $\boldsymbol{A}$ such that

$$[L(v)]_E = \boldsymbol{A}[v]_E$$

10.1.3 Basis Change

Suppose $L : \boldsymbol{V} \mapsto \boldsymbol{V}$. $E = \{v_1, \dots, v_n\}$ is a basis for $\boldsymbol{V}$, $F = \{u_1, \dots, u_n\}$ is another basis for $\boldsymbol{V}$. Then vector $u_1, \dots, u_n$ could be expressed by vectors $v_1, \dots, v_n$. So we set

$$\begin{aligned}
u_1 &= s_{11}v_1 + s_{12}v_2 + \cdots + s_{1n}v_n, \\
u_2 &= s_{21}v_1 + s_{22}v_2 + \cdots + s_{2n}v_n, \\
&\dots \\
u_n &= s_{n1}v_1 + s_{n2}v_2 + \cdots + s_{nn}v_n.
\end{aligned}$$

We could write this system into matrix form:

$$(u_1, \dots, u_n) = (v_1, \dots, v_n)\begin{pmatrix} s_{11} & s_{12} & \dots & s_{1n} \\ s_{21} & s_{22} & \dots & s_{2n} \\ \vdots & \vdots & \dots & \vdots \\ s_{n1} & s_{n2} & \dots & s_{nn} \end{pmatrix}.$$

We set $\boldsymbol{S} = (s_{ij})$. Hence we obtain:

$$(u_1, \dots, u_n) = (v_1, \dots, v_n)\boldsymbol{S}. \tag{10.1}$$

You should **prove it by yourself** that $\boldsymbol{S}$ is invertible. Hence we have:

$$(u_1, \dots, u_n)\boldsymbol{S}^{-1} = (v_1, \dots, v_n). \tag{10.2}$$

Given any vector $x \in \boldsymbol{V}$, we want to study its transformation relative to its coordinate vector. In other words, we want to study the relationship between $L(x)$ and $[x]_F$:

$$\begin{aligned}
L(x) &= \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \times [L(x)]_E \\
&= \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \times (\boldsymbol{A}[x]_E) \qquad \leftarrow \text{due to corollary (10.1)} \\
&= \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \boldsymbol{S}^{-1} \times (\boldsymbol{A}[x]_E)
\end{aligned}$$

- And we claim that $[x]_E = \boldsymbol{S}[x]_F$:
  For any vector $x \in \boldsymbol{V}$, we obtain:

  $$\begin{aligned}
  x &= \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \times [x]_E \\
  &= \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \times [x]_F \\
  &= \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \times \boldsymbol{S}[x]_F
  \end{aligned}$$

  Hence $[x]_E = \boldsymbol{S}[x]_F$.

Hence $L(x)$ could be expressed as:

$$L(x) = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \boldsymbol{S}^{-1} \times (\boldsymbol{AS}[x]_F) = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \boldsymbol{S}^{-1}\boldsymbol{AS}[x]_F$$

What do the following process mean? We know that given basis $E = \{v_1, \dots, v_n\}$, performing linear transformation on any vector x is just the same as matrix multiplication:

$$L(x) = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \times \boldsymbol{A}[x]_E$$

But what about changing another basis $F = \{u_1, \dots, u_n\}$? Do we still mutliply the same matrix $\boldsymbol{A}$? The answer is no! We must change $\boldsymbol{A}$ into $\boldsymbol{S}^{-1}\boldsymbol{AS}$:

$$L(x) = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \boldsymbol{S}^{-1}\boldsymbol{AS}[x]_F$$

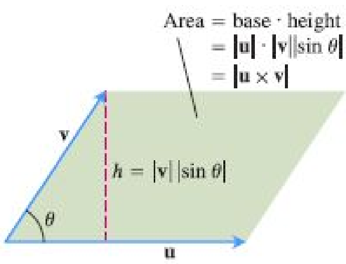
We give a definition for such phenomenon:

Definition 10.5 — Similar. Let $\boldsymbol{A}, \boldsymbol{B}$ be $n \times n$ matrix. If there exists invertible $n \times n$ matrix $\boldsymbol{S}$ such that $\boldsymbol{B} = \boldsymbol{S}^{-1}\boldsymbol{AS}$, then we say that $\boldsymbol{A}$ is **similar** to $\boldsymbol{B}$. ■

10.1.4 Determinant

The determinat of a square matrix is a single number, which contains an amazing amount of information about the matrix. It has four major uses:

- **The determinant is zero if and only if the matrix has no inverse**.
- **It can be used to calculate the area or volumn of a box:**



  For example, suppose that $\boldsymbol{u} = u_1\boldsymbol{i} + u_2\boldsymbol{j} + u_3\boldsymbol{k}, \boldsymbol{v} = v_1\boldsymbol{i} + v_2\boldsymbol{j} + v_3\boldsymbol{k}$. In order to compute the area of the parallelogram determined by $\boldsymbol{u}$ and $\boldsymbol{v}$, we just need to compute the determinant $\begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$.

- **The product of all the pivots** $= (\pm 1) \times$ **the determinant:**

  For a 2 by 2 matrix $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the pivots are a and $d - (\frac{c}{a})b$. The product of pivots is the determinant:

  $$\textbf{Product of pivots} \qquad a(d - \frac{c}{a}b) = ad - bc \qquad \textbf{which is} \quad \det \boldsymbol{A}$$

- **Compute determinants to find $\boldsymbol{A}^{-1}$ and $\boldsymbol{A}^{-1}\boldsymbol{b}$** (This formula is called **Cramer's Rule**.)

The properties of the Determinant

We don't intend to define the determinant by its formulas. It's better to start with its properties. These properties are simple, but they prepare for the formulas.

R Brackets for the matrix, straight bars for its determinant. For example,

$$\text{The determinant of } \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ is } \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

The determinant is written in two ways, $\det \boldsymbol{A}$ and $|\boldsymbol{A}|$.

We will introduce three basic properties, then we will show how properties $1-3$ derive other properties.

1. **The determinant of the n by n identity matrix is 1.**

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \qquad \text{and} \qquad \begin{vmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{vmatrix} = 1.$$

2. **The determianant changes sign when two rows are exchanged.** (sign reversal)

$$\text{Check: } \begin{vmatrix} c & d \\ a & b \end{vmatrix} = - \begin{vmatrix} a & b \\ c & d \end{vmatrix} \qquad \text{(both sides equal } bc - ad\text{).}$$

3. **The determinant is a linear function of each row separately.** (all other rows stay fixed).

$$\textbf{multiply row 1 by any number } t \qquad \begin{vmatrix} ta & tb \\ c & d \end{vmatrix} = t \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

$$\textbf{Add row 1 of } \boldsymbol{A} \textbf{ to row 1 of } \boldsymbol{B}\textbf{:} \qquad \begin{vmatrix} a_1 + a_2 & b_1 + b_2 \\ c & d \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ c & d \end{vmatrix} + \begin{vmatrix} a_2 & b_2 \\ c & d \end{vmatrix}$$

Note that this rule **deos not** mean $\det(\boldsymbol{A}+\boldsymbol{B}) = \det \boldsymbol{A} + \det \boldsymbol{B}$.
Note that this rule **does not** mean $\det(t\boldsymbol{A}) = t\det(\boldsymbol{A})$.
Actually, $\det(t\boldsymbol{A}) = t^n \det \boldsymbol{A}$. This is reasonable. Imagining that expanding a rectangle by 2, its area will increase by 4. Expand an $n-$dimensional box by t and its volumn will increase by t^n.
Pay special attention to property $1-3$. They completely determine the $\det \boldsymbol{A}$. We could stop here to find a formula for determinants. But we prefer to derive other properties that follow directly from the first three.

4. **If two rows of $\boldsymbol{A}$ are equal, then** $\det \boldsymbol{A} = \mathbf{0}$.

$$\text{Check 2 by 2: } \begin{vmatrix} a & b \\ a & b \end{vmatrix} = 0.$$

Property 4 follows from Property 2.

Proofoutline. Exchange the two equal row. The determinant $\boldsymbol{D}$ is supposed to change sign. But also the matrix is not changed, so we have $-\boldsymbol{D} = \boldsymbol{D} \implies \boldsymbol{D} = 0$. ∎

5. **Adding a constant multiple of a row to another row doesn't change** $\det \boldsymbol{A}$.

$$\begin{vmatrix} a+lc & b+ld \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} lc & ld \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + l \begin{vmatrix} c & d \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \det \boldsymbol{A}$$

Conclusion *The determinant is not changed by the usual elimination step from* $\boldsymbol{A}$ *to* $\boldsymbol{U}$. Since every row exchange reverses the sign, we have $\det \boldsymbol{A} = \pm \det \boldsymbol{U}$.

6. **If $\boldsymbol{A}$ is triangular, then** $\det \boldsymbol{A} =$ **product of diagonal entries**.

   $$\textbf{Triangular} \quad \begin{vmatrix} a & b \\ 0 & d \end{vmatrix} = ad \qquad \text{and also} \qquad \begin{vmatrix} a & 0 \\ c & d \end{vmatrix} = ad$$

   Suppose all diagonal entries of $\boldsymbol{A}$ are nonzero. We do Gaussian Elimination to convert $\boldsymbol{A}$ into diagonal matrix:

   $$\det \begin{bmatrix} a_{11} & & & \text{\Large 0} \\ & a_{22} & & \\ & & \ddots & \\ \text{\Large 0} & & & a_{nn} \end{bmatrix} = a_{11}a_{22}\dots a_{nn}.$$

   Factor a_{11} from the first row by property 3; then factor a_{22} from the second row;....... Finally the determinant is $a_{11} \times a_{22} \times a_{33} \dots \times a_{nn} \times \det \boldsymbol{I} = a_{11} \times a_{22} \times a_{33} \dots \times a_{nn}$.
7. $\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$.

   Proof.

   - If $|\boldsymbol{B}|$ is zero, it's easy to verify that $\boldsymbol{B}$ is singular, then $\boldsymbol{AB}$ is singular. Thus $\det(\boldsymbol{AB}) = 0 = \det(\boldsymbol{A})\det(\boldsymbol{B})$.
   - Suppose $|\boldsymbol{B}|$ is not zero, and $\boldsymbol{A}, \boldsymbol{B}$ is $n \times n$ matrix. Consider the ratio $D(\boldsymbol{A}) = \frac{|\boldsymbol{AB}|}{|\boldsymbol{B}|}$. *Check that this ratio has properties 1,2,3.* If so, $D(\boldsymbol{A})$ has to be the determinant, say, $|\boldsymbol{A}|$. Thus we have $|\boldsymbol{A}| = \frac{|\boldsymbol{AB}|}{\boldsymbol{B}}$:

     Property 1 *(Determinant of I)* If $\boldsymbol{A} = \boldsymbol{I}$, then the ratio becomes $D(\boldsymbol{A}) = \frac{|\boldsymbol{B}|}{|\boldsymbol{B}|} = 1$.

     Property 2 *(Sign reversal)* When two rows of $\boldsymbol{A}$ are exchanged, the same two rows of $\boldsymbol{AB}$ are also exchanged. Therefore $|\boldsymbol{AB}|$ changes sign and so does the ratio $\frac{|\boldsymbol{AB}|}{\boldsymbol{B}}$.

     Property 3 *(Linearity)* When row 1 of $\boldsymbol{A}$ is multiplied by t, so is row 1 of $\boldsymbol{AB}$. Thus the ratio is also increased by t. Thus we still have $|\boldsymbol{A}| = \frac{|\boldsymbol{AB}|}{\boldsymbol{B}}$.
     If we Add row 1 of $\boldsymbol{A}_1$ to row 1 of $\boldsymbol{A}_2$. Then row 1 of $\boldsymbol{A}_1\boldsymbol{B}$ also adds to row 1 of A_2B. By property three, determinants add. After dividing by $|\boldsymbol{B}|$, the ratios add. Hence we still have $|\boldsymbol{A}| = \frac{|\boldsymbol{AB}|}{\boldsymbol{B}}$.

     Conclusion The ratio $D(\boldsymbol{A})$ has the same three properties that defines determinant, hence it equals $|\boldsymbol{A}|$. Hence we obtain the product rule $|\boldsymbol{AB}| = |\boldsymbol{A}||\boldsymbol{B}|$.

   ■

   Immediately here follows a corollary:

   > **Corollary 10.2**
   >
   > $$\det(\boldsymbol{A}^{-1}) = \frac{1}{\det(\boldsymbol{A})}$$

8. **The transpose $\boldsymbol{A}^{\mathrm{T}}$ has the same determinant as $\boldsymbol{A}$.**

   $$\textbf{Transpose} \quad \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & c \\ b & d \end{vmatrix} \qquad \text{Both sides equal } ad - bc.$$

Proof.
- When $\boldsymbol{A}$ is singular, $\boldsymbol{A}^{\mathrm{T}}$ is also singular. Hence $|\boldsymbol{A}^{\mathrm{T}}| = |\boldsymbol{A}| = 0$.
- Otherwise $\boldsymbol{A}$ has LU decomposition $\boldsymbol{PA} = \boldsymbol{LU}$. Transposing both siders gives $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}} = \boldsymbol{U}^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}}$. By product rule we have

$$\det \boldsymbol{P} \det \boldsymbol{A} = \det \boldsymbol{L} \det \boldsymbol{U} \qquad \text{and} \qquad \det \boldsymbol{A}^{\mathrm{T}} \det \boldsymbol{P}^{\mathrm{T}} = \det \boldsymbol{U}^{\mathrm{T}} \det \boldsymbol{L}^{\mathrm{T}}.$$

  - Firstly, $\det \boldsymbol{L} = \det \boldsymbol{L}^{\mathrm{T}} = 1$. (By property 6, they both have 1's on the diagonal).
  - Secondly, $\det \boldsymbol{U} = \det \boldsymbol{U}^{\mathrm{T}}$. (By property 6, they have the same diagonal)
  - Thirdly, $\det \boldsymbol{P} = \det \boldsymbol{P}^{\mathrm{T}}$. (Verify by yourself that $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P} = \boldsymbol{I}$. Hence $|\boldsymbol{P}^{\mathrm{T}}||\boldsymbol{P}| = 1$. Since permutation matrix is obtained by exchanging rows of $\boldsymbol{I}$, the only possible value for determinant of permuation matrix is ± 1. Hence $\boldsymbol{P}$ and $\boldsymbol{P}^{\mathrm{T}}$ must both equal to 1 or both equal to -1).

  So $\boldsymbol{L}, \boldsymbol{U}, \boldsymbol{P}$ has the same determinants as $\boldsymbol{L}^{\mathrm{T}}, \boldsymbol{U}^{\mathrm{T}}, \boldsymbol{P}^{\mathrm{T}}$, Hence we have $\det \boldsymbol{A} = \det \boldsymbol{A}^{\mathrm{T}}$.

■