Linear algebra

这一次复习，我希望可以弄清楚线代的几何意义，线代究竟是怎么出来的，它究竟该怎么用，有哪些真正的物理上的含义，而不是曾经晦涩难懂的莫名其妙的定义，一些忘记了就不会再记得第二遍的东西

本次学习使用http://immersivemath.com/ila/ch00_preface/preface.html作为教材来源

https://zhuanlan.zhihu.com/p/199665495 深感认同

vector

行矢量列矢量转置The transpose of a vector

其实某种函数也可以看成矢量，举例如下：

> **Example 2.12: Polynomials up to degree 2**
> Polynomials in x up to degree 2 with real coefficients is a vector space over $\mathbb{R}$. Here if $u = u_0 + u_1 x + u_2 x^2$ and $v = v_0 + v_1 x + v_2 x^2$, where each coefficient u_i and v_i is a real number. Here vector addition $u + v$ is defined as $u + v = (u_0 + v_0) + (u_1 + v_1)x + (u_2 + v_2)x^2$ and scalar-vector multiplications is defined as $ku = ku_0 + ku_1 x + ku_2 x^2$.

基底Basis

柯西公式（等我有时间了专门来分析下柯西公式可以在哪些领域使用）

> **Popup-help:**
> The Cauchy-Schwarz inequality states that $(\mathbf{u} \cdot \mathbf{v})^2 \leq \|\mathbf{u}\|^2 \, \|\mathbf{v}\|^2$ for vectors in $\mathbb{R}^n$.

> **Popup-help:**
> The vector triple product, between three vectors, $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$, is $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$.

高斯消除法 Gaussian Elimination

> **Theorem 5.2: Gaussian Elimination Rules**
> The solutions to a linear system of equations do not change if we
>
> 1. swap the order of two equations,
> 2. multiply an equation with a constant $\neq 0$, or
> 3. add a multiple of another equation to an equation.

矩阵The Matrix

Matrices are a very powerful tool to manipulate data with.看它这里指的矩阵的意思是用来manipulate data，这个说法就很有意思，它是一种变换，是一种映射方式

把高斯消除法写成矩阵形式such as this

$$\left(\begin{array}{rrrr|r} 5 & 4 & 3 & 2 & 1496 \\ 0 & -6 & 18 & 7 & -109 \\ 0 & 0 & -30 & -65 & -2575 \\ 0 & 0 & 0 & -1395 & -40455 \end{array}\right),$$

rotation

$$\begin{aligned} \mathbf{q} &= \begin{pmatrix} r\cos(\theta+\phi) \\ r\sin(\theta+\phi) \end{pmatrix} = \begin{pmatrix} r(\cos\theta\cos\phi - \sin\theta\sin\phi) \\ r(\sin\theta\cos\phi + \cos\theta\sin\phi) \end{pmatrix} \\ &= \underbrace{\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}}_{\mathbf{R}(\phi)} \underbrace{\begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}}_{\mathbf{p}}, \end{aligned}$$

scaling 缩放

Definition 6.11: Two-Dimensional Scaling Matrix
A scaling matrix is defined by

$$\mathbf{S}(f_x, f_y) = \begin{pmatrix} f_x & 0 \\ 0 & f_y \end{pmatrix},$$

where f_x is the factor that is applied in the x-dimension and f_y is applied to the y-dimension.

求矩阵的逆矩阵的方法

AX=IB=》IX=A^-1B

$$\left(\begin{array}{rrr|rrr} 1 & 0 & 0 & 4 & -13 & -6 \\ 0 & 1 & 0 & -7 & 24 & 11 \\ 0 & 0 & 1 & 2 & -7 & -3 \end{array}\right).$$

和高斯规则相同

正交矩阵Orthogonal Matrices

Definition 6.18: Orthogonal Matrix

An orthogonal matrix B is a square matrix where the column vectors constitute an orthonormal basis.正交矩阵B是一个正方形矩阵，其中列向量构成了正交标准。

> **Theorem 6.11: Orthogonal Matrix Equivalence**
> The following are equivalent
> (i) The matrix $\mathbf{B}$ is orthogonal.
> (ii) The column vectors of $\mathbf{B}$ constitute an orthonormal basis.
> (iii) The row vectors of $\mathbf{B}$ constitute an orthonormal basis.
> (iv) $\mathbf{B}^{-1} = \mathbf{B}^{\mathrm{T}}$

an orthonormal basis：它的列向量都是单位向量，且两两正交

旋转矩阵

$$= \underbrace{\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}}_{\mathbf{R}(\phi)}$$ 或

$$\begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

是正交矩阵

正交矩阵对矢量进行处理不会改变它的模长

正交矩阵乘正交矩阵后还是正交矩阵

正交矩阵处理后的矢量之间的性质不变

$$(\mathbf{Bu}) \cdot (\mathbf{Bv}) = \mathbf{u} \cdot \mathbf{v},$$

正交矩阵相当于坐标变换（找基底什么的）认真看下面这个例子就能理解了

Example 6.15: Change of Base using Orthogonal Matrices
Assume we have two orthonormal bases, $\{\mathbf{e}_1, \mathbf{e}_2\}$ and $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2\}$, defined as

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \text{ and } \hat{\mathbf{e}}_1 = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}, \hat{\mathbf{e}}_2 = \begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}. \tag{6.96}$$

It is easy to check that $\|\mathbf{e}_i\| = 1$ and $\|\hat{\mathbf{e}}_i\| = 1$ for $i \in \{1, 2\}$ and that $\mathbf{e}_1 \cdot \mathbf{e}_2 = 0$ and $\hat{\mathbf{e}}_1 \cdot \hat{\mathbf{e}}_2 = 0$, i.e., we have two orthonormal bases, per Definition 3.3. Now, it is possible to use Theorem 6.10 to find out what the matrices look like that expresses these bases. However, an alternative way when dealing with orthonormal bases is simply to imagine the vectors $(1, 0)$ and $(0, 1)$ expressed in $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2\}$ and then figure out how to set up a matrix that transforms those vectors into $\{\mathbf{e}_1, \mathbf{e}_2\}$. It is rather simple as seen below.

$$\underbrace{\begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}}_{\mathbf{A}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix} \text{ and } \underbrace{\begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}}_{\mathbf{A}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix} \tag{6.97}$$

Now assume that we have a vector $\mathbf{v} = (3/4, 1/2)$ in $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2\}$ and we want to transform that vector into $\{\mathbf{e}_1, \mathbf{e}_2\}$. It is simply a matter of multiplying with the matrix $\mathbf{A}$ above, i.e.,

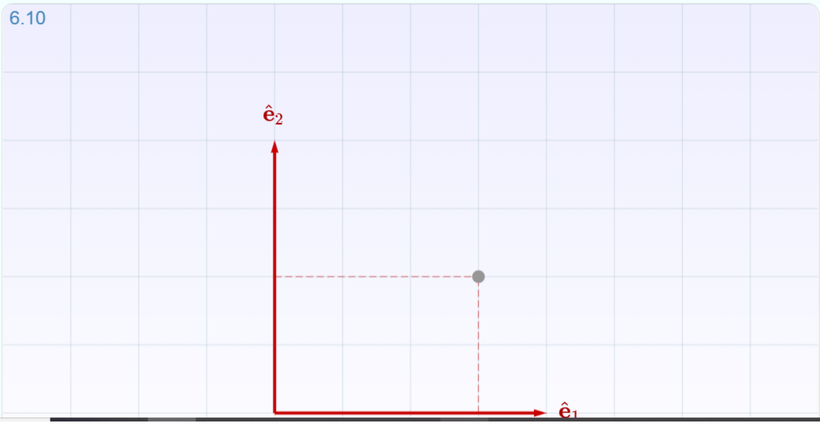
$$\mathbf{A}\mathbf{v} = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} \frac{3}{4} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{3\sqrt{3}-2}{8} \\ \frac{2\sqrt{3}+3}{8} \end{pmatrix} \tag{6.98}$$

So $\mathbf{A}$ can be used to transform a vector in $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2\}$ so that it instead is expressed in $\{\mathbf{e}_1, \mathbf{e}_2\}$. This indicates that $\mathbf{A}^{\mathrm{T}}$ can be used to transform a vector in $\{\mathbf{e}_1, \mathbf{e}_2\}$ so that instead is expressed in $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2\}$.

Now assume we have yet another orthonormal basis, $\{\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2\}$, whose corresponding transform matrix is $\mathbf{B}$. This means that to take a vectors from $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2\}$ to $\{\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2\}$, you first use $\mathbf{A}$ to get to $\{\mathbf{e}_1, \mathbf{e}_2\}$ and then $\mathbf{B}^{\mathrm{T}}$ to get to $\{\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2\}$. If we were to apply this to a vector, $\mathbf{v}$, then would be expressed as

$$\mathbf{v}' = \mathbf{B}^{\mathrm{T}}\mathbf{A}\mathbf{v}, \tag{6.99}$$

where $\mathbf{v}'$ is expressed in $\{\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2\}$. The first steps of this are visualized in Interactive Illustration 6.10.



行列式Determinants

Definition 7.1: Determinant
The determinant det is a scalar function of a square matrix $\mathbf{A}$ fulfilling the following three properties:

$$\begin{array}{lll} (i) & \det(\mathbf{I}) = 1 & \text{(determinant of unit matrix)} \\ (ii) & |\ldots \ \mathbf{a}_i \ \ldots \ \mathbf{a}_i \ \ldots| = 0 & \text{(zero if two columns are equal)} \\ (iii) & |\ldots \ \lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 \ \ldots| = \lambda_1|\ldots \ \mathbf{a}_1 \ \ldots| + \lambda_2|\ldots \ \mathbf{a}_2 \ \ldots| & \text{(linear in each column)} \end{array}$$

Theorem 7.1: Properties of the determinant

$$\begin{array}{lll} (iv) & |\ldots \ \mathbf{0} \ \ldots| = 0 & \text{(zero if one column is zero)} \\ (v) & |\ldots \ \mathbf{a}_i \ \ldots \ \mathbf{a}_j \ \ldots| = -|\ldots \ \mathbf{a}_j \ \ldots \ \mathbf{a}_i \ \ldots| & \text{(swapping columns)} \\ (vi) & |\ldots \ \mathbf{a}_i \ \ldots \ \mathbf{a}_j \ \ldots| = |\ldots \ \mathbf{a}_i + \lambda\mathbf{a}_j \ \ldots \ \mathbf{a}_j \ \ldots| & \text{(adding to another column)} \\ (vii) & \det(\mathbf{A}) = \det(\mathbf{A}^{\mathrm{T}}) & \text{(transpose)} \\ (viii) & \det(\mathbf{A}\mathbf{B}) = \det(\mathbf{A})\det(\mathbf{B}) & \text{(product)} \\ (ix) & \det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})} & \text{(inverse)} \end{array}$$

伴随矩阵

Definition 7.9: Adjoint matrix
The adjoint matrix $\operatorname{adj}(\mathbf{A})$ is a matrix whose elements at position (i, j) is $(-1)^{i+j} D_{ji}$, where similar to the previous section D_{ij} is the determinant of the matrix A after removing row i and column j.

Theorem 7.8:
If $\mathbf{A}$ is a square matrix and $\operatorname{adj}(\mathbf{A})$ is its adjoint matrix, then

$$\mathbf{A}(\operatorname{adj}(\mathbf{A})) = (\operatorname{adj}(\mathbf{A}))\mathbf{A} = \det(\mathbf{A})\mathbf{I}.$$

If $\det(\mathbf{A}) \neq 0$ then $\mathbf{A}$ is invertible and the inverse is given by

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \operatorname{adj} \mathbf{A}.$$

求方程的方法2

Theorem 7.9:
If $\mathbf{A}$ is a square matrix and $\det \mathbf{A} \neq 0$ then the solution $\mathbf{x}$ to the matrix equation $\mathbf{Ax} = \mathbf{y}$ has

$$x_i = \frac{\det(\mathbf{a}_1 \quad \ldots \quad \mathbf{a}_{i-1} \quad \mathbf{y} \quad \mathbf{a}_{i+1} \quad \ldots \quad \mathbf{a}_n)}{\det \mathbf{A}}.$$

Cramer's rule.

秩 Rank

线性空间 linear vector space

1. The zero vector $\mathbf{0} \in W$.
2. If $\mathbf{v}_1 \in W$ and $\mathbf{v}_2 \in W$ then $\mathbf{v}_1 + \mathbf{v}_2 \in W$.
3. If $\mathbf{v} \in W$ and $\lambda \in F$ then $\lambda\mathbf{v} \in W$.

Let $\mathbf{A}$ be an $m \times n$ matrix. Let V be the linear vector space $\mathbb{R}^m$. Let $W \subset V$ be the subset of vectors $\mathbf{v} \in V$ that can be generated as $\mathbf{v} = \mathbf{Au}$ for some $\mathbf{u} \in \mathbb{R}^n$. Then W is a linear vector space. This is true since

Definition 8.1: Null Space
If $\mathbf{A}$ is an $m \times n$ matrix (i.e., with m rows and n columns), then the entire set of solutions to $\mathbf{Ax} = \mathbf{0}$ is called *the null space* of $\mathbf{A}$.

Definition 8.3: Null Vectors
Let $\mathbf{p}_1, \ldots, \mathbf{p}_k$ be a basis for the null space. A solution, $\mathbf{x}$, is on the form $\mathbf{x}(t_1, \ldots t_k) = t_1\mathbf{p}_1 + \cdots + t_k\mathbf{p}_k$, where k is the dimension of the null space. The basis vectors $\mathbf{p}_i$ are sometimes called *null vectors*, although sometimes any vector in the null space is called a null vector.

nullity，是Null Space 的维度，是方程未知量个数减A的秩

Theorem 8.8: Dimension Theorem
For an $m \times n$ matrix $\mathbf{A}$, i.e., with n columns, it holds that

$$\operatorname{rank}(\mathbf{A}) + \operatorname{nullity}(\mathbf{A}) = n. \quad (8.32)$$

秩

Definition 8.6: Row and Column Rank
Let $\mathbf{A}$ be a matrix. The *row rank* of $\mathbf{A}$, denoted $\operatorname{rowrank}(\mathbf{A})$, is the maximum number of linearly independent row vectors of $\mathbf{A}$. Similarly, the *column rank*, denoted $\operatorname{colrank}(\mathbf{A})$, is the maximum number of linearly independent column vectors of $\mathbf{A}$.

Theorem 8.2:
Neither of the operations of Gaussian elimination (Theorem 5.2) changes the row space of an $m \times n$ matrix $\mathbf{A}$ after applying the operation.

Theorem 8.6: Rank of Product
The rank of a product $\mathbf{A} = \mathbf{BC}$ is less than or equal to the rank of the terms, i.e.,

$$\operatorname{rank} \mathbf{A} \leq \operatorname{rank} \mathbf{C},$$

and

$$\operatorname{rank} \mathbf{A} \leq \operatorname{rank} \mathbf{B}.$$

Theorem 8.7:
Let $\mathbf{x}_{\mathrm{h}}$ be the solution to $\mathbf{Ax} = \mathbf{0}$ and let $\mathbf{x}_{\mathrm{p}}$ be the solution to a particular system of equations $\mathbf{Ax} = \mathbf{y}$, in which case the "entire" solution to $\mathbf{Ax} = \mathbf{y}$ is

$$\mathbf{x}_{\mathrm{tot}} = \mathbf{x}_{\mathrm{p}} + \mathbf{x}_{\mathrm{h}}. \tag{8.30}$$

linear mappings

In linear algebra, the term *mapping* is traditionally used instead of function, but the meaning is the same,

Definition 9.2: Domain, Codomain, and Range of a Mapping
Assume we have a mapping $y = F(x)$ where $x \in N$ and $y \in M$. Then N is the *domain* of the mapping, and M is the *codomain* of the mapping. The *range* (or alternatively, the *image*) of the mapping is the set V_F, where

$$V_F = \{F(x) | x \in N\}.$$

It is now easy to see that the matrix $\mathbf{A}$ is just a two-dimensional rotation matrix as defined in Definition 6.10 in Chapter 6. When a mapping can be written in matrix form, i.e., in the form $\mathbf{y} = \mathbf{Ax}$, we call $\mathbf{A}$ the *transformation matrix.*

Example 9.2: Shopping Cart to Cost
Assume that a shop only sells packages of penne, jars of Arrabiata sauce, and bars of chocolate. The contents of your shopping cart can be modelled as a vector space. Introduce addition of two shopping carts as putting all of the items of both carts in one cart. Introduce multiplication of a scalar as multiplying the number of items in a shopping cart with that scalar. Notice that here, there are practical problems with multiplying a shopping cart with non-integer numbers or negative numbers, which makes the model less useful in practice. Introduce a set of basis shopping carts. Let $\mathbf{e}_1$ correspond to the shopping cart containing one package of penne, let $\mathbf{e}_2$ correspond to the shopping cart containing one jar of Arrabiata sauce, and let $\mathbf{e}_3$ correspond to the shopping cart containing one bar of chocolate. Then each shopping cart $\mathbf{x}$ can be described by three coordinates (x_1, x_2, x_3) such that $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$.

There is a mapping from shopping carts $\mathbf{x}$ to price $y \in \mathbb{R}$. We can introduce the matrix $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \end{pmatrix}$ where a_{11} is the price of a package of penne, a_{12} is the price of a jar of Arrabiata sauce and a_{13} is the price of a bar of chocolate. The price y can now be expressed as $y = \mathbf{Ax}$.

In real life this map is often non-linear, e.g., a shop might have campaigns saying 'buy 3 for the price of 2'. But modelling the mapping as a linear map is often a reasonable and useful model. Again (as is common with mathematical modelling) there is a discrepancy between mathematical model and reality. The results of mathematical analysis must always be used with reason and critical thinking. Even if the cost of a shopping cart of 1 package of penne is 10, it does not always mean that you can sell packages of penne to the store for 10 each.

Theorem 9.1: Matrix Form of Linear Mappings
A mapping $\mathbf{y} = F(\mathbf{x})$ can be written in the form $\mathbf{y} = \mathbf{Ax}$ (matrix form) if and only if it is linear.

这个证明有点意思，贴上来

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2,$$
$$\mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2. \tag{9.12}$$

Inserting the expression for $\mathbf{x}$ in $\mathbf{y} = F(x)$, we get

$$\mathbf{y} = F(\mathbf{x}) = F(x_1\mathbf{e}_1 + x_2\mathbf{e}_2), \tag{9.13}$$

and since F is linear, we can apply the first and second conditions of linearity,

$$\mathbf{y} = F(x_1\mathbf{e}_1) + F(x_2\mathbf{e}_2) = x_1F(\mathbf{e}_1) + x_2F(\mathbf{e}_2). \tag{9.14}$$

Since F maps one vector to another vector, $F(\mathbf{e}_1)$ must also be a vector that can be expressed in the basis. Assume it has the coordinates $\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$ in the base $\mathbf{e}_1$, $\mathbf{e}_2$,

$$F(\mathbf{e}_1) = a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2. \tag{9.15}$$

Likewise, we assume

$$F(\mathbf{e}_2) = a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2. \tag{9.16}$$

We can now continue the expansion of $F(\mathbf{x})$ as

$$\begin{gathered}\mathbf{y} = x_1(a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2) + x_2(a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2) = \\ (x_1a_{11} + x_2a_{12})\mathbf{e}_1 + (x_1a_{21} + x_2a_{22})\mathbf{e}_2\end{gathered} \tag{9.17}$$

Comparing this expression to the second row of Equation (9.12), we understand that y_1 must equal $a_{11}x_1 + a_{12}x_2$ and $y_2 = a_{21}x_1 + a_{22}x_2$. We have

$$\boxed{\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}} \tag{9.18}$$

that is

$$\mathbf{y} = \mathbf{A}\mathbf{x} \tag{9.19}$$

Theorem 9.2: Mapping of Basis Vectors
For a linear mapping $\mathbf{y} = F(\mathbf{x})$ written on the form $\mathbf{y} = \mathbf{A}\mathbf{x}$ in the base $\mathbf{e}_1,\ \mathbf{e}_2,\ \ldots,\ \mathbf{e}_n$, the column vectors of $\mathbf{A}$ are the images of the basis vectors, $\mathbf{a}_{,1} = F(\mathbf{e}_1),\ \mathbf{a}_{,2} = F(\mathbf{e}_2),\ \ldots,\ \mathbf{a}_{,n} = F(\mathbf{e}_n)$.

Example 9.3: Finding a Linear Mapping's Matrix
A linear mapping $\mathbf{y} = F(\mathbf{x})$ rotates a two-dimensional vector $\mathbf{x}$ counterclockwise 90 degrees. Find the transformation matrix $\mathbf{A}$ of the matrix form $\mathbf{y} = \mathbf{A}\mathbf{x}$ when the standard orthonormal basis $\mathbf{e}_1 = (1,0)$, $\mathbf{e}_2 = (0,1)$ is used.

The first column vector of $\mathbf{A}$ will be the image of the first basis vector $\mathbf{e}_1$, which is the x-axis. Rotating that 90 degrees counterclockwise brings it parallel with the y-axis, which has coordinates $(0,1)$. Thus the first column vector is $\mathbf{a}_{,1} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

The second column vector will be the image of the second basis vector, which is the y-axis. Rotating the y-axis 90 degrees counter clockwise makes it ending up in $(-1,0)$. The second column vector is thus $\mathbf{a}_{,1} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ and we can write $\mathbf{A}$ as

$$\mathbf{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{9.25}$$

This is shown in Interactive Illustration 9.7.

这个还是挺有意思的，F里面每一个列向量都是对ei进行F变换得来的

Theorem 9.3: Composition of Linear Mappings
If $\mathbf{y} = F(\mathbf{x})$ and $\mathbf{z} = G(\mathbf{y})$ are both linear mappings then the composite mapping $\mathbf{z} = G(F(\mathbf{x}))$ is also linear.

好，开始矩阵相乘的运用了

Definition 9.4: Injective mappings
A mapping $y = F(x)$ is *injective*, if any two different vectors $\mathbf{x}_1 \neq \mathbf{x}_2$ will always give rise to two different images $F(\mathbf{x}_1) \neq F(\mathbf{x}_2)$.

Another way to state this is to say that an injective mapping has the property that if two images $F(\mathbf{x}_1)$ and $F(\mathbf{x}_2)$ are equal, then $\mathbf{x}_1$ must equal $\mathbf{x}_2$.

——对应

Definition 9.5: Surjective mapping
A mapping $\mathbf{y} = F(\mathbf{x})$, where $\mathbf{x} \in N$ and $\mathbf{y} \in M$ is *surjective* if the range V_F is equal to to the codomain M, $V_F = M$.

全体映射

Definition 9.6: Bijective mapping
A mapping is *bijective* if it is both injective and surjective.

——对应+全体映射

A-》B B-》A 好了应该要开始可逆了

Theorem 9.4: Equivalence of Inverse Mapping
For a linear mapping $\mathbf{y} = F(\mathbf{x})$ from $\mathbf{x} \in \mathbb{R}^n$ to $\mathbf{y} \in \mathbb{R}^n$, the following three statements are equivalent:

i. The mapping F is bijective.
ii. The transformation matrix for F is invertible.
iii. The images $F(\mathbf{e}_1), \ldots, F(\mathbf{e}_n)$ of the basis vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ constitute a basis in $\mathbb{R}^n$.

Theorem 9.5: Inverse Mapping Matrix
For a bijective linear mapping $\mathbf{y} = F(\mathbf{x})$ with transformation matrix $\mathbf{A}$, the inverse mapping $\mathbf{x} = F^{-1}(\mathbf{y})$ is linear and has the transformation matrix $\mathbf{A}^{-1}$.

越来越有意思咯，这种mapping

特征值和特征向量 Eigenvalues and Eigenvectors

Definition 10.1: Eigenvalue and Eigenvector
Let $\mathbf{A}$ be a square matrix. A non-zero column vector $\mathbf{v}$ is called an eigenvector if $\mathbf{A}\mathbf{v}$ is parallell to $\mathbf{v}$, i.e., if there exists a scalar λ such that

$$\mathbf{A}\mathbf{v} = \lambda \mathbf{v}. \tag{10.5}$$

The scalar λ is then called an eigenvalue.

Definition 10.2: Characteristic Polynomial
Let $\mathbf{A}$ be a square matrix of size $n \times n$. The polynomial

$$p_{\mathbf{A}}(\lambda) = \det(\lambda I - \mathbf{A}) \tag{10.6}$$

is then called the *characteristic polynomial*.

Theorem 10.2:
The characteristic polynomial is a polynomial of degree n in λ. The eigenvalues λ to the matrix $\mathbf{A}$ are the roots of the characteristic polynomial $p_{\mathbf{A}}(\lambda)$.

特征值方程

Definition 10.3:
A linear mapping F is diagonalizable if there is a basis, such that the transformation matrix is a diagonal matrix.

Definition 10.4:
A matrix $\mathbf{A}$ is diagonalizable, if there exists an invertible matrix $\mathbf{V}$ and a diagonal matrix $\mathbf{D}$ such that

$$\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1}. \tag{10.18}$$

对角化

Theorem 10.3:
A matrix $\mathbf{A}$ is diagonalizable if and only if there exists a basis of eigenvectors. Furthermore, in the factorization $\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1}$, the columns of $\mathbf{V}$ correspond to eigenvectors and the corresponding elements of the diagonal matrix $\mathbf{D}$ are the eigenvalues.

Theorem 10.4:
If the $n \times n$ matrix $\mathbf{A}$ has n different eigenvalues, then $\mathbf{A}$ is diagonalizable. Furthermore, n eigenvectors corresponding to different eigenvalues are always linearly independent.

Definition 10.5: Algebraic Multiplicity, Eigenspace, and Geometric Multiplicity
Let λ be an eigenvalue to the matrix $\mathbf{A}$. The multiplicity of λ with respect to the characteristic polynomial is called the algebraic multiplicity of λ. The nullspace of $(\lambda I - \mathbf{A})$ is called the eigenspace of the matrix $\mathbf{A}$ associated with the eigenvalue λ. The geometric multiplicity of λ is the dimension of the nullspace of $(\lambda I - \mathbf{A})$, i.e., the dimension of the eigenspace.

Theorem 10.5:
The geometric multiplicity of an eigenvalue λ is always less than or equal to the algebraic multiplicity.

对那个多项式的解释

Theorem 10.6:
A matrix $\mathbf{A}$ is diagonalizable if and only if the geometric and algebraic multiplicity is the same for each eigenvalue. or equivalently for an $n \times n$ matrix $\mathbf{A}$, if the sum of the dimensions of the eigenspaces is n, then $\mathbf{A}$ is diagonalizable.

Definition 10.6: Complex Eigenvalue and Eigenvector
Let $\mathbf{A}$ be a square matrix with real or complex entries. A non-zero column vector $\mathbf{v}$ with complex entries is called an eigenvector if $\mathbf{Av}$ is parallell to $\mathbf{v}$, i.e., if there exists a complex scalar λ, such that

$$\mathbf{Av} = \lambda \mathbf{v}. \tag{10.52}$$

The scalar λ is then called an eigenvalue.

A useful property of symmetric matrices is summarized in the following theorem.

Theorem 10.7:
The eigenvalues of a symmetric matrix $\mathbf{A}$ are all real.

Theorem 10.8:
For symmetric matrix $\mathbf{A}$, it is possible to find an orthogonal matrix $\mathbf{U}$ and diagonal matrix $\mathbf{D}$ such that

$$\mathbf{A} = \mathbf{UDU}^{-1} = \mathbf{UDU}^{T}. \tag{10.55}$$

In other words, it is possible to diagonalise using an orthonormal basis of eigenvectors.

这个网站对这些东西的定义没有学校给它定义定义的好，还是看学校的那个对角化知识把，反正现在也在教嘛，可以教到了再细看

结尾的话

其实把线代看下来，会真的对国内的线代教学产生恶感，流畅的逻辑就应该是从向量到线性空间，从线性空间的变换到矩阵，从矩阵到行列式再到mapping，再从mapping到特征值特征向量这些东西

现在感觉最深的就是线代的使用，尤其是增广矩阵的运算和可逆的运算，要常学

其实本来以为9-5号就能学完的，居然硬生生学了三天，还是记忆不行呀