Instituto de Física
Universidade de São Paulo

Notas de Álgebra Linear

Victor Hugo dos Santos LINS
victorlins@usp.br

What has been affirmed without proof
can also be denied without proof.
– Euclid.

Combinações lineares, espaço gerado e vetores de base.

Considere um vetor arbitrário $(3, -2)$. Podemos interpretar os números 3 e -2 como *escalares*, isto é, números que serão multiplicados por outros vetores para modificar a sua escala. A situação mais comum é que esses escalares multipliquem os vetores chamados *standard*: em física, costumamos chamá-los de $\hat{\imath}$ e $\hat{\jmath}$. Vetorialmente, podemos escrevê-los como

$$\hat{\imath} = (1, 0), \quad \hat{\jmath} = (0, 1)$$

Portanto, o vetor $(3, -2)$ nada mais é do que

$$(3, -2) = 3\hat{\imath} - 2\hat{\jmath} = 3 \cdot (1, 0) + (-2) \cdot (0, 1)$$

Esses vetores *standard* possuem um nome especial: são os *vetores base* do sistema de coordenadas xy, ou espaço vetorial $\mathbb{R}^2$.

Na prática, sempre que escrevemos um vetor, como $(3, -2)$, estamos escolhendo implicitamente um conjunto de vetores base, e a escolha natural é a dos vetores *standard*. O que ocorre é que a expressão para um vetor pode ser expandida como a multiplicação dos escalares pelos respectivos vetores da base.
Qualquer vetor no plano $\mathbb{R}^2$ pode ser obtido através de um par diferente de vetores base, não precisa ser *standard*. Apenas utilizamos estes vetores em particular como base porque eles são mais conveniente para nós fazermos cálculos. Ainda, reconheça que podemos generalizar essa conclusão para espaços vetoriais de dimensão superior a $\mathbb{R}^2$. O que quero dizer é que dados quaisquer dois vetores não colineares no $\mathbb{R}^2$ nós podemos obter qualquer outro vetor do $\mathbb{R}^2$ através de uma *combinação linear* desses dois vetores.

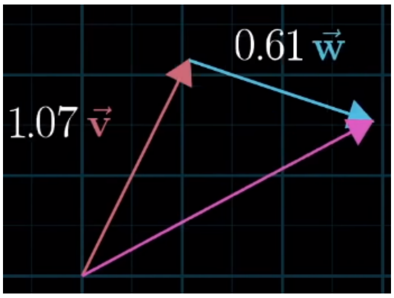


Imagem: 3Blue1Brown.

Uma *combinação linear* nada mais é do que o somatório dos vetores base multiplicados respectivamente por escalares arbitrários. Para exemplificar, sejam $\vec{v}$ e $\vec{w}$ dois vetores arbitrários, e α e β dois escalares pertencentes a $\mathbb{R}$, também arbitrários. Dizemos que uma combinação linear de $\vec{v}$ e $\vec{w}$ é

$$\alpha\vec{v} + \beta\vec{w}$$

Note que qualquer vetor do $\mathbb{R}^2$ pode ser obtido a partir dessa combinação linear — basta que você escolha α e β convenientemente.
Se dois vetores escolhidos forem colineares, isto é, possuem uma mesma direção, então dizemos que os vetores são *linearmente dependentes*. Caso contrário, chamamos de *linearmente independentes.* Ainda, podemos definir o conceito de *espaço gerado*, que nada mais é do que o conjunto de todas as combinações lineares de um conjunto de vetores. Por exemplo, se dois vetores são linearmente independentes, o espaço gerado por eles é justamente $\mathbb{R}^2$, uma vez que todas as combinações lineares de dois vetores linearmente independentes pode definir qualquer ponto do plano. Por outro lado, se dois vetores são linearmente dependentes, o espaço gerado por eles nada mais é do que $\mathbb{R}$ (uma reta), ou ainda um simples ponto na origem, que ocorre quando ambos são vetores nulos.

Transformações Lineares.

Uma transformação linear nada mais é do que uma função, isso é, uma coisa que recebe um argumento, realiza algum tipo de operação, e nos retorna uma outra coisa. A razão para que haja esse nome "transformação" é porque normalmente os argumentos que essa função receberá serão *vetores*, e a interpretação geométrica de aplicar uma transformação linear em um vetor é transformá-lo em um outro vetor. Vamos chamar essa transformação de ϕ.

Dizemos que ϕ é linear se, e somente se,

$$\phi(\alpha\vec{v} + \beta\vec{w}) = \alpha\phi(\vec{v}) + \beta\phi(\vec{w})$$

Isto é, fazer uma transformação linear em um vetor que já foi escalado é o mesmo que escalar o vetor imagem da transformação aplicada no vetor original, e também a transformação de uma soma de vetores é a soma da transformação de cada um dos vetores individualmente. Se essas condições forem satisfeitas, então podemos dizer que a transformação ϕ é linear.

Estamos interessados em saber como descrever essas transformações lineares aplicadas em vetores. Uma forma interessante de tratar isso no $\mathbb{R}^2$ é checar onde os vetores base $\hat{\imath}$ e $\hat{\jmath}$ "aterrissam" após a transformação ter sido concluída. Na prática, o que estou querendo dizer é isso:

Suponha que você deseja saber o resultado de uma transformação linear ϕ aplicada no vetor $\vec{v} = (3, -2)$. Vamos escrever $\vec{v}$ em função dos vetores base *standard*:

$$\vec{v} = 3\hat{\imath} - 2\hat{\jmath} = 3(1,0) - 2(0,1)$$

Agora, vamos aplicar ϕ a $\vec{v}$:

$$\phi(\vec{v}) = \phi(3(1,0) - 2(0,1)) = 3\phi(1,0) - 2\phi(0,1)$$

Onde o último passo só foi possível porque estamos assumindo que ϕ é linear. Percebe como o efeito de aplicar ϕ a $\vec{v}$ se resume a entender onde $\hat{\imath}$ e $\hat{\jmath}$ aterrissam quando submetidos a transformação linear ϕ? A própria transformação fica em função de $\phi(\hat{\imath}) = \phi(1,0)$ e $\phi(\hat{\jmath}) = \phi(0,1)$ — é por isso, sacou? Nós não precisamos modificar nada nos escalares!

Portanto, a conclusão que nós chegamos é que toda a informação necessária para descrever a aplicação ϕ em qualquer vetor de $\mathbb{R}^2$ é saber onde os vetores $\hat{\imath}$ e $\hat{\jmath}$ aterrissam. Armazenamos essas "posições de aterrissagem" para o caso de dimensão 2 em uma matrix 2×2, e dizemos que essa matriz descreve a aplicação ϕ. Simbolicamente:

$$A_\phi \equiv \begin{pmatrix} \alpha & \beta \\ \gamma & \theta \end{pmatrix}$$

onde (α, β) é onde o vetor $\hat{\imath}$ aterrissa após a transformação linear escrito na base *standard*, isto é, $(\alpha, \beta) = \alpha\hat{\imath} + \beta\hat{\jmath}$, e (γ, θ) é o análogo para $\hat{\jmath}$.

O que isso nos diz é que dado um vetor qualquer, como $\vec{v} = (3, -2)$, para saber qual novo vetor será obtido após a atuação da transformação linear ϕ basta que façamos

$$\vec{v} \cdot A_\phi = \begin{pmatrix} 3 & -2 \end{pmatrix} \cdot \begin{pmatrix} \alpha & \beta \\ \gamma & \theta \end{pmatrix} = 3 \cdot \begin{pmatrix} \alpha & \beta \end{pmatrix} - 2 \cdot \begin{pmatrix} \gamma & \theta \end{pmatrix} = \begin{pmatrix} 3\alpha - 2\gamma & 3\beta - 2\theta \end{pmatrix}$$

Note que nessa nova descrição $\vec{v}_1 = (\alpha, \beta)$ e $\vec{v}_2 = (\gamma, \theta)$ estão funcionando como nossos novos vetores base, uma vez que estamos descrevendo o vetor $(3, -2)$ como uma combinação linear desses vetores base, cujos coeficientes da combinação são 3 e -2, respectivamente, isto é $(3, -2) = 3 \cdot \vec{v}_1 - 2 \cdot \vec{v}_2$. Basta passar um tempo olhando para a expressão acima. Por essa razão, dizemos que A_ϕ funcionou como uma *matriz mudança de base* — o efeito resultante da transformação linear foi o de modificar a base com que estamos trabalhando, coisa que iremos conversar em mais detalhes em breve.

Vamos agora tratar de um exemplo para não ficar tanto nas abstrações.

Exemplo. Consideremos a transformação linear de rotação em 90° no sentido anti-horário. Isso significa que se essa transformação for aplicada em um vetor, o efeito resultante vai ser o de girá-lo 90° no sentido anti-horário.

Então, seja $\vec{v} = (a, b) = a(1,0) + b(0,1)$ um vetor qualquer, descrito na base *standard*. Como discutido anteriormente, a transformação será totalmente descrita se tivermos o lugar onde cada vetor base aterrissa. Após um giro de 90° no sentido anti-horário, o vetor $\hat{\imath} = (1,0)$ vai para $(0,1)$, enquanto que o vetor $\hat{\jmath} = (0,1)$ vai para $(-1,0)$. Portanto, a matriz que descreve a transformação linear A_ϕ envolvida é

$$A_\phi = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Com efeito, se quisermos saber qual é o vetor resultante da transformação linear ϕ aplicada no vetor $\vec{v}$ definido anteriormente, fazemos

$$v \cdot A_\phi = (a, b) \cdot \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = (-b, a)$$

Uma observação relevante: se a matriz 2×2 que descreve ϕ tiver um conjunto de vetores *linearmente dependentes*, o efeito resultante é o de que os vetores base são transformados em dois vetores que possuem mesma direção, então os vetores que antes descreviam juntos um plano inteiro agora descrevem apenas uma linha.

Composição de Transformações Lineares.

Comumente estamos interessados no efeito resultante de executar uma transformação ϕ e logo a seguir uma transformação ψ. Existe uma maneira de descrever o resultado dessas duas transformações de uma forma única? A resposta breve é sim, e de forma bem simples. Suponha que você possui um vetor $\vec{v} = (a, b)$ arbitrário. Ao realizar uma transformação ϕ nesse vetor, sabemos que obtemos um novo vetor $\vec{w}$ tal que

$$\vec{w} = \vec{v} \cdot A_\phi$$

Agora, se desejamos aplicar uma transformação ψ logo após ter aplicado ϕ, basta que apliquemos ψ em $\vec{w}$, isto é, obtenhamos um novo vetor $\vec{r}$ tal que

$$\vec{r} = \vec{w} \cdot A_\psi = \vec{v} \cdot A_\phi \cdot A_\psi$$

por transitividade, temos

$$\vec{r} = \vec{v} \cdot (A_\phi \cdot A_\psi) = \vec{v} \cdot A_{\phi\psi}$$

onde $A_{\phi\psi}$ é a composição das transformações lineares A_ϕ seguida de A_ψ. Aplicar $A_{\phi\psi}$ a um vetor é o mesmo que aplicar A_ϕ e depois aplicar A_ψ. Então, a composição de transformações lineares é equivalente a multiplicar as matrizes que as definem.

$$A_{\phi\psi} = A_\phi \cdot A_\psi$$

O DETERMINANTE.

Uma vez que compreendemos o conceito de transformação linear, uma pergunta interessante a ser feita é a seguinte:

Suponha que nós conhecemos a área de uma certa região do plano englobada pelo paralelogramo formado pelos vetores da base *standard*, por exemplo um quadrado de lado 1. Se eu aplico uma transformação linear arbitrária ϕ, como podemos prever qual será a nova área daquela região que anteriormente tratava-se do quadrado de lado 1? Afinal, sabemos que após a aplicação de ϕ, em geral, os vetores da base irão aterrissar em posições diferentes das iniciais, portanto a área englobada pelo paralelogramo será também diferente. De que forma as áreas descritas por figuras no plano mudam após a transformação linear?

Para compreender a resposta para essa pergunta, comece pensando em uma aplicação linear ϕ específica:

$$A_\phi = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$$

O que essa transformação faz com os vetores da base é levar $\hat{i} = (1,0)$ para $(3,0)$ e levar $\hat{j} = (0,1)$ para $(0,2)$. Observe que o efeito resultante em $\hat{i}$ foi o de reescalar $\hat{i}$ em 3 vezes (esse é um fato relevante que faremos uso depois).

Pense na área descrita pelo paralelogramo formado anteriormente pelos dois vetores base *standard*, trata-se da área de um quadrado de lado 1. Após a aplicação da transformação linear ϕ, os vetores da nova base descrevem um paralelogramo equivalente a um retângulo de lados 2 e 3. Portanto, após a atuação de ϕ, a área $A = 1$ do quadrado de lado 1 foi reescalada para a área $A = 6$ de um retângulo de lados 2 e 3, ou seja, foi reescalada por um fator igual a 6. Reconheça que qualquer outra área arbitrária do plano $\mathbb{R}^2$ seria igualmente escalada pelo mesmo fator 6, uma vez que qualquer área pode ser descrita em função de quadradinhos de lado 1 a partir dos vetores da base. Aqui vem o pulo do gato: esse fator especial que nós obtemos "6" é chamado de *determinante* da nossa transformação linear. Ou seja, quando calculamos o determinante de A_ϕ nós estamos prevendo qual é o fator que irá escalar qualquer área arbitrária em $\mathbb{R}^2$ após a atuação de ϕ. Essa é uma interpretação geométrica belíssima para o determinante, e nos auxiliará de forma substancial em muitos procedimentos que virão no futuro.

Note que o determinante, em geral, pode ser positivo, negativo ou nulo. No caso em que ele é positivo, ocorre uma simples mudança de escala nas áreas por um fator igual ao determinante. Caso ele seja negativo, o módulo do determinante será o fator com que as áreas são reescaladas, mas pelo fato do determinante em si ser negativo o efeito que isso causa na região que descreve a área é uma *troca de orientação*. O que quero dizer com isso é que se você adotava um certo vetor normal a essa região plana orientado "para cima", após a atuação de ϕ além da área ser reescalada por $\det(\phi)$, o vetor normal passará a apontar "para baixo".

Por fim, se $\det(\phi) = 0$ isso significa que após a atuação de ϕ nós transformamos qualquer área arbitrária em zero, cuja interpretação geométrica significa transformar um plano em uma reta (que possui área nula) ou ainda, em um caso mais extremo, num ponto. Esse último caso é extremamente importante, porque sempre que estivermos diante de uma transformação linear com determinante nulo, isso quer dizer que a atuação dessa aplicação irá reduzir o espaço em pelo menos uma dimensão. Essa consequência é fundamental em Álgebra Linear, como nós veremos em próximos problemas.

Por fim, precisamos aprender a calcular os determinantes. Aqui eu ensino a fazê-lo para matrizes de transformações lineares 2×2 e também 3×3, reconhecendo que aprender a fazer contas em qualquer ordem $n \times n$ pode ser feito em qualquer livro ou ainda fazendo uma rápida pesquisa no Google — o intuito aqui é focar em aprender o conteúdo.

1. Determinante de matrizes 2×2

   Seja A uma matriz arbitrária

   $$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

   O determinante de A é dado por

   $$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

   Exemplo 1. Calcule

   $$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$$

   Solução. Por definição,

   $$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 1 \cdot 4 - 2 \cdot 3 = 4 - 6 = -2.$$

2. Determinante de matrizes 3×3

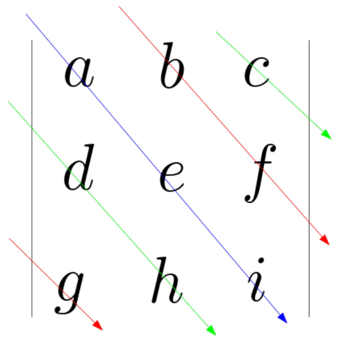
   Seja C uma matriz arbitrária

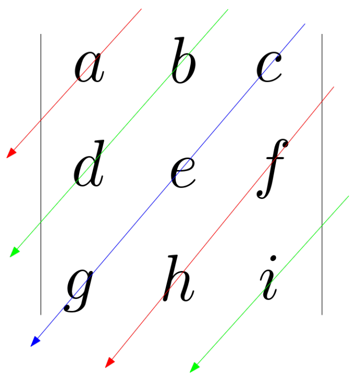
   $$C = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

   O determinante de C é dado por

$$\det(C) = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + chd - ceg - fha - idb$$

A forma como eu gosto de pensar para realizar o cálculo do determinante de matrizes 3×3 é a seguinte: eu sei que o cálculo se resume a multiplicar 3 números por vez, somando alguns e subtraindo os outros. Os números que irei somar são os que obtenho através da multiplicação guiada por linhas decrescentes da esquerda para a direita, enquanto que os números que irei subtrair são os que obtenho através da multiplicação guiada por linhas decrescentes da direita para a esquerda. Observe a imagem a seguir para entender o que eu quero dizer com essas linhas decrescentes.





Na primeira figura, eu sempre começo pela diagonal principal, onda a reta decrescente me guia para multiplicar aei. Depois eu prossigo para a próxima diagonal, onde a reta decrescente me guia para multiplicar bf. Está faltando um termo no produto, porque nós sabemos que os produtos são três a três. Eu preciso atravessar para o outro lado da diagonal, e se eu tentar passar pela diagonal onde tem h e d eu vou acabar multiplicando quatro termos, o que está incorreto. Isso me faz concluir que eu preciso passar pela diagonal onde está o g, o que resulta no produto bfg. Por fim, eu multiplico os três termos que faltam, seguindo a diagonal de c e logo após isso a diagonal que contém h e d, resultando em chd. Portanto, eu acabei de obter as três triplas que serão somadas na primeira parte do cálculo do determinante, utilizando uma lógica bem simples e que eu acho bem mais preferível do que perder tempo redesenhando as primeiras duas colunas. Uma vez que você se adeque a este novo raciocínio, funcionará de forma automática e infalível.

A seguir, aplicaremos a mesma lógica porém agora para as linhas decrescentes da direita para a esquerda, que irão nos guiar para as triplas que serão subtraídas. Você deve ser capaz de intuir o procedimento baseando no que fizemos no parágrafo anterior, e o resultado final nos levará a concluir que

$$\det(C) = aei + bfg + chd - ceg - fha - idb$$

Exemplo 2. Calcule

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$$

Solução. Por definição,

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 1 \cdot 5 \cdot 9 + 2 \cdot 6 \cdot 7 + 3 \cdot 8 \cdot 4 - 3 \cdot 5 \cdot 7 - 6 \cdot 8 \cdot 1 - 9 \cdot 4 \cdot 2 = 0$$

Uma consequência muito importante da forma como definimos o determinante, sua interpretação geométrica e o seu cálculo, é que nós somos capazes de afirmar se um conjunto de vetores é linearmente dependente ou não simplesmente calculando o determinante da matriz cujas linhas são os vetores, vamos justificar isso.

Suponha que você possui uma tripla de vetores $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ linearmente dependentes. Escrevendo eles em uma matriz, temos

$$A_\phi = \begin{pmatrix} \vec{v}_1 \\ \vec{v}_2 \\ \vec{v}_3 \end{pmatrix}$$

Utilizando a interpretação de transformações lineares, onde uma matriz pode ser compreendida como uma transformação linear que transforma os vetores standard nos vetores relativos a cada linha da matriz da transformação, estamos diante de uma aplicação A_ϕ que transforma três vetores linearmente independentes (os vetores base *standard*) em três vetores linearmente dependentes (a tripla $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$). Portanto, antes possuíamos três vetores capazes de gerar um espaço $\mathbb{R}^3$ e agora possuímos três vetores que no máximo conseguem descrever um plano $\mathbb{R}^2$: *houve uma redução na dimensionalidade do espaço gerado!* Isso significa que a nossa matriz de aplicação deve ter determinante nulo, argumento esse justificado pelo desenvolvimento da interpretação geométrica do determinante feita nas páginas anteriores. Logo, dado qualquer conjunto de vetores, se desejamos saber se eles são linearmente dependentes basta que nós escrevamos uma matriz onde cada linha é relativa a cada vetor e calculamos o determinante dessa matriz — se for nulo, então os vetores são linearmente dependentes; se for diferente de zero, então eles são linearmente independentes.

MUDANÇAS DE BASE.

Em geral, sabemos que ao escrever um vetor arbitrário $\vec{v} = (a, b)$ estamos escolhendo implicitamente um par de vetores base, o que quase sempre será os vetores *standard* por conveniência. Entretanto, uma competência necessária para trabalhar com a Álgebra Linear é saber efetuar mudanças de base. Por vezes, receberemos transformações lineares que na base *standard* será algo complicado e aleatório, mas se formos capazes de realizar uma mudança de base talvez possamos trabalhar com essa transformação em um formato muito mais simples. É essa competência que iremos desenvolver a seguir.

Suponha que nós estamos adotando, inicialmente, o par de vetores *standard*, isto é,

$$\vec{v}_1 = (1, 0) \qquad \vec{v}_2 = (0, 1)$$

Suponha também que exista um segundo par de vetores base, dados por

$$\vec{b}_1 = (2, 1) \qquad \vec{b}_2 = (-1, 1)$$

Comece reparando que esses vetores da segunda base estão *escritos na nossa base standard*. Quero dizer, por exemplo,

$$\vec{b}_1 = (2, 1) = 2\hat{\imath} + 1\hat{\jmath} = 2(1, 0) + 1(0, 1)$$

No entanto, na perspectiva de quem possui essa base $\{\vec{b}_1, \vec{b}_2\}$, esses vetores na verdade são os vetores $(1, 0)$ e $(0, 1)$, porém escritos na base $\{\vec{b}_1, \vec{b}_2\}$. Em outras palavras, $\vec{b}_1 = (1, 0)$ na base $\{\vec{b}_1, \vec{b}_2\}$, mas $\vec{b}_1 = (2, 1)$ na base $\{\vec{v}_1, \vec{v}_2\}$. Portanto, geometricamente falando, um vetor que para nós parece $\vec{v} = (a, b)$ na base *standard*, pode ser algo totalmente diferente na outra base. Por essa razão, é importante conseguir traduzir entre diferentes bases.

Exemplo. Suponha que alguém que possui como vetores base $\vec{b}_1 = (2, 1)$ e $\vec{b}_2 = (-1, 1)$, escritos na base *standard*, afirma estar visualizando o vetor $\vec{x} = (-1, 2)$, isto é, na base dela. Como obter o vetor $\vec{x}$ na base *standard*?

Solução. Quando a pessoa afirma estar vendo o vetor $\vec{x} = (-1, 2)$ é necessário recordar-se que essa visualização é utilizando os vetores base $\{\vec{b}_1, \vec{b}_2\}$, ou seja

$$(-1, 2) \equiv -1\vec{b}_1 + 2\vec{b}_2 = -1(2, 1) + 2(-1, 1) \equiv (-1, 2) \cdot \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} = (-1, 2) \cdot \begin{pmatrix} \vec{b}_1 \\ \vec{b}_2 \end{pmatrix}$$

Ou seja, para que nós possamos entender qual é o vetor que a pessoa afirma estar vendo ser $\vec{x} = (-1, 2)$, nós precisamos multiplicar $(-1, 2)$ por uma matriz cujas linhas são os vetores da base dessa pessoa escritos na nossa base *standard*. Em outras palavras, essa matriz é justamente a transformação linear que transforma os nossos vetores *standard* nos vetores de base dessa pessoa — lembra da noção de transformação e aterrissagem? Portanto, a matriz cujas linhas são os vetores da base da outra pessoa funcionará como *matriz mudança de base*. Assim,

$$\vec{x} = (-1, 2) \cdot \begin{pmatrix} \vec{b}_1 \\ \vec{b}_2 \end{pmatrix} = (-1, 2) \cdot \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} = -1(2, 1) + 2(-1, 1) = (-4, 1)$$

É o vetor $(-1, 2)$ que a outra pessoa afirma estar vendo, porém *dessa vez ele está escrito na nossa base*, isto é

$$\vec{x} \equiv -1\vec{b}_1 + 2\vec{b}_2 \equiv -4\vec{v}_1 + 1\vec{v}_2$$

onde $\vec{v}_1 = (1, 0)$ e $\vec{v}_2 = (0, 1)$. Além disso, a matriz mudança de base é

$$C = \begin{pmatrix} \vec{b}_1 \\ \vec{b}_2 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}$$

Na prática, o que essa matriz faz quando aplicada a um vetor arbitrário (a, b) escrito na base da outra pessoa é transformar os nossos vetores base *standard* nos vetores base da outra pessoa, escritos na nossa base, cuja implicação numérica é descrever o vetor que a outra pessoa entende por (a, b) na base dela em uma outra numeração (c, d) que representa o mesmo vetor no espaço porém na nossa base. **Leia esse parágrafo pelo menos umas 3 ou 4 vezes**.

Nós poderíamos ter empregado, desde o início, o raciocínio contrário, bastando que para isso em vez de empregarmos como matriz mudança de base a matriz que transforma os nossos vetores base nos vetores base da outra pessoa, a matriz que transforma os vetores base da outra pessoa nos nossos vetores base, isto é, C^{-1}. Dessa forma, se tivéssemos algum vetor, por exemplo, $\vec{y} = (1, 2)$ escrito nos nossos vetores base, bastaria que fizéssemos $\vec{y} \cdot C^{-1}$ para encontrar o vetor equivalente ao nosso $(1, 2)$ porém na base da outra pessoa. É assim que funciona para transformarmos um vetor de uma base para outra e vice-versa.

Entretanto, ainda há um problema não resolvido: como fazer para converter uma transformação linear de uma base para outra? Isto é, se possuímos uma matriz de transformação linear que nos diz onde os nossos vetores base *standard* vão aterrissar, como que a outra pessoa descreveria essa mesma transformação porém na base dela?

Suponha um vetor $\vec{v} = (a, b)$, escrito na base da outra pessoa. Suponha, também, que a transformação linear ϕ, escrita com nossos vetores base, seja

$$A_\phi = \begin{pmatrix} \alpha & \beta \\ \gamma & \theta \end{pmatrix}$$

Então, para que a outra pessoa possa saber como seria a transformação ϕ escrita na nossa base aplicada ao vetor $\vec{v} = (a, b)$ escrito na base dela, o primeiro passo seria transformar os nossos vetores base *standard* nos vetores base dela, escritos na nossa base. Para tanto, sendo a base dela $\{\vec{b}_1, \vec{b}_2\}$, temos

$$\vec{v} \cdot \begin{pmatrix} \vec{b}_1 \\ \vec{b}_2 \end{pmatrix}$$

Nesse momento nós transformamos os nossos vetores base nos vetores base dela, escritos na nossa base. Porém, como esses vetores bases dela ainda estão escritos nos nossos vetores base, podemos aplicar a nossa transformação ϕ a esse vetor e calcular onde os vetores base dela escritos na nossa base irão aterrissar após a transformação, isto é

$$\vec{v} \cdot \begin{pmatrix} \vec{b}_1 \\ \vec{b}_2 \end{pmatrix} \cdot \begin{pmatrix} \alpha & \beta \\ \gamma & \theta \end{pmatrix}$$

Por fim, agora que nós já possuímos o resultado da transformação aplicada ao vetor $\vec{v}$ na base da outra pessoa escrita nos vetores da nossa base *standard*, iremos trazer de volta os vetores base da outra pessoa escritos na nossa base para os nossos vetores base escritos na nossa base, multiplicando o resultado atual pela matriz mudança de base inversa, isto é

$$\vec{v} \cdot \begin{pmatrix} \vec{b}_1 \\ \vec{b}_2 \end{pmatrix} \cdot \begin{pmatrix} \alpha & \beta \\ \gamma & \theta \end{pmatrix} \cdot \begin{pmatrix} \vec{b}_1 \\ \vec{b}_2 \end{pmatrix}^{-1}$$

Portanto, o efeito resultante foi de ter aplicado ao vetor $\vec{v}$ na base da outra pessoa a transformação

$$C \cdot A_\phi \cdot C^{-1}$$

onde

$$C = \begin{pmatrix} \vec{b}_1 \\ \vec{b}_2 \end{pmatrix}$$

$$A_\phi = \begin{pmatrix} \alpha & \beta \\ \gamma & \theta \end{pmatrix}$$

Assim, o procedimento resultante para mudar a base de uma transformação linear A_ϕ é o de realizar o produto

$$CA_\phi C^{-1}$$

onde C é a matriz mudança de base, constituída de forma que suas linhas são os vetores da nova base.

AUTOVETORES E AUTOVALORES.

Entre as diversas transformações lineares existentes, existem algumas que possuem um efeito especial quando aplicadas em um vetor: o efeito resultante após a atuação da transformação linear foi simplesmente o de *reescalar* o vetor. Simbolicamente, se a matriz da transformação linear é dada por A_ϕ e o vetor arbitrário em que estamos aplicando a transformação é $\vec{v}$, então

$$\vec{v} \cdot A_\phi = \lambda \vec{v}$$

onde $\lambda \in \mathbb{R}$ é um escalar arbitrário.

Reconheça, também, a existência de uma matriz de transformação linear cujo efeito resultante quando aplicada em um vetor é o de *não fazer nada*, isto é, mantém os vetores da base *standard* em seus respectivos lugares. Chamamos a essa matriz que não faz nada de matriz identidade, e escrevemos identidade 2×2 como

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Portanto, retornando a nossa equação para a aplicação de uma transformação linear num vetor $\vec{v}$ tal que o efeito resultante é o de simplesmente reescalá-lo, temos

$$\vec{v} \cdot A_\phi = \lambda \vec{v} = \lambda I \vec{v}$$

onde a última igualdade foi possível porque aplicar I a v é o mesmo que não ter feito nada na prática. Entretanto, para nós, esse I estar nesse lugar no nosso cálculo é de extrema importância, porque agora podemos subtrair de ambos os lados $\lambda \cdot I\vec{v}$ e obter

$$\vec{v} \cdot A_\phi - \lambda \cdot I\vec{v} = \vec{0}$$

e colocando $\vec{v}$ em evidência, temos

$$\vec{v} \cdot \big(A_\phi - \lambda I\big) = \vec{0}$$

Estamos interessados nos vetores $\vec{v}$ que resolvem esta equação, mas não nos interessa o caso óbvio/trivial em que $\vec{v} = (0, 0, 0)$. Queremos obter os outros $\vec{v}$ que resolvem essa equação. Para isso, reconheça que o que essa equação nos está dizendo é que quando a matriz $(A_\phi - \lambda I)$ é aplicada no vetor $\vec{v}$ o resultado é o vetor nulo: *houve uma redução de dimensionalidade no espaço gerado*! Isso implica que o determinante dessa matriz deve ser nulo, isto é,

$$\det\big(A_\phi - \lambda I\big) = 0$$

Essa última equação é chamada de *equação característica*, e o polinômio em λ que obtemos quando tentamos resolvê-la é chamado de *polinômio característico.*

As soluções dessa equação são justamente os valores de λ que tornam a equação característica verdade, e quando substituímos os valores de λ na equação somos plenamente capazes de descobrir quais são os vetores respectivos $\vec{v}$ que tornam a equação verdade. Para cada valor de λ temos pelo menos um vetor $\vec{v}$ associado. Essas soluções λ da equação nós chamamos de *autovalores* de A_ϕ e aos vetores $\vec{v}$ associados nós chamamos de *autovetores* de A_ϕ.

Na prática, esses autovetores de A_ϕ são os vetores que, uma vez aplicado A_ϕ neles, o efeito resultante é o de apenas mudarmos sua escala, e o fator em que é escalado é justamente o autovalor λ associado.

A utilidade desses autovalores e autovetores é principalmente a de que podemos fatorar a matriz A_ϕ como uma matriz diagonal através da mudança de base da matriz A_ϕ para uma base cujos vetores são os autovetores de A_ϕ. Vale lembrar que queremos muito que A_ϕ esteja escrita numa base em que ela fica diagonal, uma vez que matrizes diagonais são excelentes para fazer cálculos rapidamente e também possuem a qualidade de que comutam com qualquer outra matriz. Portanto, dada uma matriz A_ϕ, nós resolvemos a equação característica para ela, obtemos seus autovalores, seus autovetores, e escrevemos os autovetores numa nova matriz C, que servirá como nossa *matriz mudança de base*. Realizando a operação justificada na seção anterior

$$CA_\phi C^{-1}$$

Iremos obter A_ϕ numa base em que ela fica diagonal.