Revista Eletrônica Matemática e Estatística em Foco
Volume 11 - Número 1
Outubro 2024

Artigo de Divulgação

A Matriz de Gram, a Regra de Cramer e a Distância entre Subespaços Afins

Daniel Cariello
Instituto de Matemática e Estatística - Universidade Federal de Uberlândia
dcariello@gmail.com

Resumo

A distância entre dois subespaços afins normalmente é obtida através do cálculo da projeção de um vetor a um certo subespaço vetorial. Entretanto isso pode ser evitado utilizando a regra de Cramer e o determinante da matriz de Gram. Nesse pequeno artigo de divulgação mostramos como essas duas ferramentas combinadas fornecem uma fórmula para distância entre subespaços afins. As famosas fórmulas da distância de ponto a reta no $\mathbb{R}^2$ e de ponto a plano no $\mathbb{R}^3$ são casos particulares dessa fórmula.

Palavras-chaves: Matriz de Gram, Regra de Cramer, Subespaços afins.

Abstract

The distance between two affine subspaces is usually obtained by calculating the projection of a certain vector onto a certain subspace. This can be avoided by using the Cramer's rule and the determinant of a certain Gram matrix. In this letter we show how to combine these two tools to obtain a formula for the distance between two affine subspaces. The famous formulas for the distance from point to line in $\mathbb{R}^2$ and for the distance from point to plane in $\mathbb{R}^3$ are instances of this more general formula.

Keywords: Cramer's rule, Gram matrix, Affine subspaces.

1 Introdução

Todos os alunos que já estudaram geometria analítica, por exemplo na referência (WEXLER, 1962), conhecem a famosa fórmula da distância do ponto (x_0, y_0) a reta $ax + by + c = 0$,

$$\frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}.$$

Uma maneira simples de obtê-la, que evita o uso de vetores, é através do seguinte argumento.

Seja $C = (x_0, y_0)$ e considere a reta $ax + by + c = 0$. A partir de C ande na vertical até encontrar a reta no ponto $B = (x_0, y_1)$. Novamente a partir de C ande na horizontal até encontrar a reta no ponto $A = (x_1, y_0)$. Isso te dá um triângulo retângulo conforme figura abaixo.

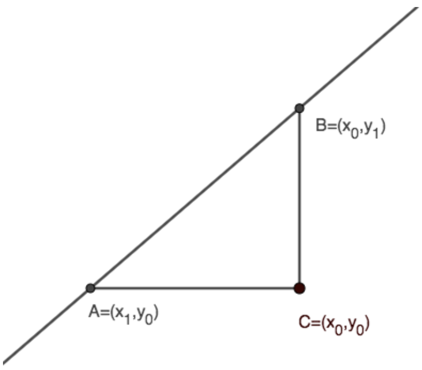

Figura 1: Distância de ponto a reta no $\mathbb{R}^2$
Fonte: Próprio autor

Podemos calcular a área desse triângulo ABC de duas maneiras diferentes.

- Primeiro usando como base o lado AC e obtemos como área
$$\frac{|AC||BC|}{2} = \frac{|x_1 - x_0||y_1 - y_0|}{2}.$$

- Segunda maneira usando como base o lado AB e obtemos como área
$$\frac{|AB|h}{2} = \frac{\sqrt{|y_1 - y_0|^2 + |x_1 - x_0|^2}h}{2},$$
onde o h é a distância de C a reta $ax + by + c = 0$ (a quantidade que estamos interessados).

Igualando as duas fórmulas das áreas temos
$$|y_1 - y_0||x_1 - x_0| = \sqrt{|y_1 - y_0|^2 + |x_1 - x_0|^2}h, \tag{1}$$

Como $A = (x_1, y_0)$ e $B = (x_0, y_1)$ pertencem a $ax + by + c = 0$ então
$$\begin{cases} ax_1 + by_0 + c = 0, \text{ ou seja, } x_1 = \dfrac{-c - by_0}{a}. \\ ax_0 + by_1 + c = 0, \text{ ou seja, } y_1 = \dfrac{-c - ax_0}{b}. \end{cases}$$

Assim,
$$|y_1 - y_0| = \frac{|ax_0 + by_0 + c|}{|b|} \quad \text{e} \quad |x_1 - x_0| = \frac{|ax_0 + by_0 + c|}{|a|}$$

Substituindo essas informações na equação (1) obtemos $h = \dfrac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}$.

De forma muito parecida podemos obter a outra famosa fórmula da distância do ponto (x_0, y_0, x_0) ao plano $ax + by + cz + d = 0$ no $\mathbb{R}^3$,
$$\frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

Seja $D = (x_0, y_0, z_0)$ e considere o plano $ax+by+cz+d = 0$. A partir de D ande na direção do eixo z até encontrar o plano no ponto $A = (x_0, y_0, z_1)$. Novamente a partir de D ande na direção do eixo y até encontrar o plano no ponto $B = (x_0, y_1, z_0)$. Finalmente a partir de D ande na direção do eixo x até encontrar o plano no ponto $C = (x_1, y_0, z_0)$. Isso te dá uma pirâmide conforme figura abaixo.

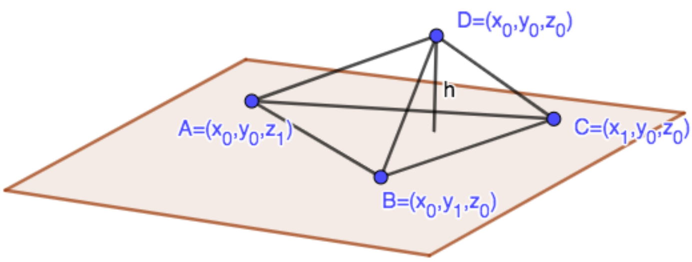

Figura 2: Distância de ponto a plano no $\mathbb{R}^3$
Fonte: Próprio autor

Como a partir de D andamos na direção dos eixos x, y e z então os ângulos da pirâmide formados no vértice D são todos ângulos retos.

Podemos calcular o volume dessa pirâmide $ABCD$ de duas maneiras diferentes.

- Primeiro usando como base o lado BDC e altura AD. Obtemos como volume
$$\frac{1}{3}\frac{|BD||CD|}{2}|AD| = \frac{|x_1 - x_0||y_1 - y_0||z_1 - z_0|}{6}.$$

- Segunda maneira usando como base o lado ABC, cuja área vale metade do módulo do produto vetorial $B - A \times C - A$, e altura h que é a desejada distância de D ao plano $ax + by + cz + d = 0$. Assim obtemos como volume
$$\frac{1}{3}\frac{|B - A \times C - A|}{2}h = \frac{1}{6}\sqrt{|y_1 - y_0|^2|z_1 - z_0|^2 + |x_1 - x_0|^2|y_1 - y_0|^2 + |x_1 - x_0|^2|z_1 - z_0|^2}h.$$

Igualando as duas fórmulas do volume obtemos

$$|y_1 - y_0||x_1 - x_0||z_1 - z_0| = \sqrt{|y_1 - y_0|^2|z_1 - z_0|^2 + |x_1 - x_0|^2|y_1 - y_0|^2 + x_1 - x_0|^2|z_1 - z_0|^2}h. \tag{2}$$

Como $A = (x_0, y_0, z_1)$, $B = (x_0, y_1, z_0)$ e $C = (x_1, y_0, z_0)$ pertencem a $ax + by + cz + d = 0$, obtemos

$$z_1 = \frac{-d - ax_0 - by_0}{c}, \quad y_1 = \frac{-d - ax_0 - cz_0}{b} \ \text{ e } \ x_1 = \frac{-d - by_0 - cz_0}{a}.$$

Assim
$$\begin{cases} |z_1 - z_0| = \dfrac{|ax_0 + by_0 + cz_0 + d|}{|c|} \\[2ex] |y_1 - y_0| = \dfrac{|ax_0 + by_0 + cz_0 + d|}{|b|} \\[2ex] |x_1 - x_0| = \dfrac{|ax_0 + by_0 + cz_0 + d|}{|a|} \end{cases}$$

Substituindo essas informações na equação (2) obtemos

$$h = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

Talvez o leitor esteja incomodado com a validade dessa demonstração quando a ou b ou c valem 0. Realmente o argumento nesses casos não funciona. Essa não é a única limitação desse argumento, por exemplo, esse argumento não funciona no $\mathbb{R}^n$, com $n > 3$, simplesmente porque um plano no $\mathbb{R}^n$, com $n > 3$, não tem essa equação $ax + by + cz + d = 0$. Calcular distância de ponto a plano no $\mathbb{R}^n$ requer outros métodos. Além disso, por que nos restringir simplesmente as distâncias entre pontos, retas e planos?

Esse pequeno artigo de divulgação trata exatamente da distância entre dois subespaços afins em espaços vetoriais reais com produto interno e dimensão finita, dos quais são exemplos os pontos, as retas e os planos. Veremos que existe uma fórmula envolvendo determinantes que fornece a distância entre dois subespaços afins (Veja Corolário 4.1) e que só precisamos de um velho conhecido de Álgebra Linear para obtê-la: a regra de Cramer.

Começaremos a próxima seção falando sobre subespaços afins e reduziremos o problema de calcular a distância entre dois deles ao problema de calcular a distância de um vetor a um subespaço vetorial (Veja Observação 2.1).

Na seção 3, veremos como calcular a distância de vetor a subespaço utilizando a projeção no subespaço (Veja Proposição 3.2), ou seja, se calcularmos a projeção já somos capazes de achar a distância de vetor a subespaço.

A projeção está completamente relacionada com a matriz de Gram. A matriz de Gram dos vetores $w_1, \ldots, w_n$ de um espaço vetorial com produto interno é definida como a matriz de ordem n cuja entrada ij é ocupada pelo número $\langle w_i, w_j \rangle$. Denotaremos essa matriz por $Gram(w_1, \ldots, w_n)$.

Entretanto, nós não estamos interessados em calcular a projeção de um vetor a um subespaço, mas sim na distância dele ao subespaço. Existe uma maneira de calcular essa distância sem calcular a projeção. A distância de um vetor v a um subespaço W, com base $w_1, \ldots, w_n$, é dada pela fórmula

$$\sqrt{\frac{\det(Gram(w_1, \ldots, w_n, v))}{\det(Gram(w_1, \ldots, w_n))}}.$$

Provamos essa fórmula na seção 4 utilizando a regra de Cramer (Veja Teorema 4.1).

Por fim na seção 5, mencionaremos algumas consequências dessa fórmula. Em particular, veremos que as fórmulas de distância de ponto a reta no $\mathbb{R}^2$ e de distância de ponto a plano no $\mathbb{R}^3$ são casos particulares dessa fórmula.

Nós aprendemos essa fórmula na demonstração do Teorema de Müntz na referência (CAROTHERS, 2009). Metade desse teorema é consequência dessa fórmula. Descrevemos essa metade no fim desse artigo. Recomendamos a leitura desse teorema e da referência (CAROTHERS, 2009) inteira, pois são muito interessantes.

2 Distância entre subespaços afins

Vamos começar com a definição de subespaço afim e depois definiremos a distância entre dois deles. Veremos nessa pequena seção que a distância entre dois deles se reduz a distância de um vetor a subespaço.

Definição 2.1

Seja V um espaço vetorial, W um subespaço de V e $a \in V$. Um subespaço afim é o conjunto do tipo $W + a = \{w + a, w \in W\}$, ou seja, ele é a translação de um subespaço.

Exemplo 2.1

1. Um vetor a é um subespaço afim $W + a$, onde $W = \{\vec{0}\}$,
2. Uma reta é um subespaço afim $W + a$, onde W é o subespaço de dimensão 1,
3. Um plano é um subespaço afim $W + a$, onde W é o subespaço de dimensão 2.

Definição 2.2

Seja V um espaço vetorial real com produto interno, onde $W_1 + a$ e $W_2 + b$ são dois subespaços afins de V. Seja $|v| = \sqrt{\langle v, v\rangle}$ para todo $v \in V$. Definimos a distância entre $W_1 + a$ e $W_2 + b$ por

$$\inf_{w_1 \in W_1,\ w_2 \in W_2} |(w_1 + a) - (w_2 + b)| = \inf_{w_1 \in W_1,\ w_2 \in W_2} |(a - b) - (w_2 - w_1)|.$$

Observação 2.1

Note que todo vetor do subespaço $W_1 + W_2 = \{r + s, r \in W_1 \text{ e } s \in W_2\}$ também pode ser escrito como $w_2 - w_1$, onde $w_2 = s$ e $w_1 = -r$. Isso significa que a distância entre $W_1 + a$ e $W_2 + b$ pode ser escrita como

$$\inf_{v \in W_1 + W_2} |(a - b) - v|.$$

Isso significa que a distância entre os subespaços afins $W_1 + a$ e $W_2 + b$ é a distância do vetor $a - b$ ao subespaço $W_1 + W_2$.

Conclusão: Para calcular distância entre subespaços afins basta saber como calcular distância entre vetor e subespaço vetorial. Iremos nos concentrar nisso nas próximas seções.

3 Projeção de vetor a subespaço e Matriz de Gram

Seja V um espaço vetorial real com produto interno e W um subespaço de V de dimensão finita. Veremos nessa seção que o vetor de W mais próximo do vetor v é a projeção de v em W que é definida da seguinte maneira.

Definição 3.1

Definimos a projeção de v em W, denotada por $proj_W(v)$, como sendo o vetor com as seguintes propriedades:

(i) $proj_W(v) \in W$,

(ii) $v - proj_W(v)$ é ortogonal a todo $w \in W$. Em outras palavras, $v - proj_W(v) \in W^{\perp}$, onde $W^{\perp}$ é o complemento ortogonal de W.

Vejamos abaixo que realmente existe um vetor com essas propriedades.

Proposição 3.1

Seja V um espaço vetorial real com produto interno e W um subespaço de V de dimensão finita. As seguintes afirmações são válidas:

(i) A $proj_W(v)$ existe e é única.

(ii) Se a dimensão de V é finita então $proj_{W^{\perp}}(v) = v - proj_W(v)$ e $|v - proj_W(v)|^2 + |v - proj_{W^{\perp}}(v)|^2 = |v|^2$.

Demonstração: Seja $w_1, \ldots, w_n$ uma base de W. Pela Definição 3.1, queremos que $proj_W(v) = a_1 w_1 + \ldots + a_n w_n$ e que $v - proj_W(v)$ seja ortogonal a todo $w \in W$, mas para isso ocorrer basta que $v - proj_W(v)$ seja ortogonal a base.

Assim $\langle w_i, v - a_1 w_1 - \ldots - a_n w_n \rangle = 0$ para todo $i = 1, \ldots, n$, ou seja,

$$\begin{gathered} \langle w_1, v \rangle = a_1 \langle w_1, w_1 \rangle + \ldots + a_n \langle w_1, w_n \rangle \\ \vdots \\ \langle w_n, v \rangle = a_1 \langle w_n, w_1 \rangle + \ldots + a_n \langle w_n, w_n \rangle. \end{gathered}$$

Podemos reescrever essas equações no seguinte sistema linear.

$$\begin{pmatrix} \langle w_1, w_1 \rangle & \ldots & \langle w_1, w_n \rangle \\ \vdots & \ddots & \vdots \\ \langle w_n, w_1 \rangle & \ldots & \langle w_n, w_n \rangle \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \langle w_1, v \rangle \\ \vdots \\ \langle w_n, v \rangle \end{pmatrix}. \tag{3}$$

Vimos na introdução desse artigo que a matriz desse sistema é a matriz de Gram dos vetores $w_1, \ldots, w_n$ (denotada por $Gram(w_1, \ldots, w_n)$). Para que $proj_W(v)$ exista e seja única, basta que a matriz $Gram(w_1, \ldots, w_n)$ seja invertível. Dessa maneira os valores de $a_1, \ldots, a_n$ serão únicos e

$$proj_W(v) = a_1 w_1 + \ldots + a_n w_n.$$

Mas isso é fácil de ver. Suponha que $Gram(w_1, \ldots, w_n)$ tenha um vetor $(b_1, \ldots, b_n) \neq (0, \ldots, 0)$ no núcleo. Assim

$$\begin{pmatrix} \langle w_1, w_1 \rangle & \ldots & \langle w_1, w_n \rangle \\ \vdots & \ddots & \vdots \\ \langle w_n, w_1 \rangle & \ldots & \langle w_n, w_n \rangle \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Mas isso implica que

$$(b_1, \ldots, b_n) \begin{pmatrix} \langle w_1, w_1 \rangle & \ldots & \langle w_1, w_n \rangle \\ \vdots & \ddots & \vdots \\ \langle w_n, w_1 \rangle & \ldots & \langle w_n, w_n \rangle \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = 0.$$

O lado esquerdo da igualdade acima pode ser reescrito como

$$\langle b_1 w_1 + \ldots + b_n w_n, b_1 w_1 + \ldots + b_n w_n \rangle = 0,$$

ou seja, $b_1 w_1 + \ldots + b_n w_n = \vec{0}$, o que é absurdo pois nem todos os b_i são nulos e $w_1 \ldots, w_n$ são linearmente independentes.

Portanto o núcleo de $Gram(w_1, \ldots, w_n)$ só contém o vetor nulo, isso garante que ela é invertível. Vejamos agora a segunda parte.

Se V tiver dimensão finita então acabamos de ver que $proj_{W^\perp}(v)$ existe e é o único vetor de V que satisfaz

- $proj_{W^\perp}(v) \in W^\perp$,
- $v - proj_{W^\perp}(v)$ é ortogonal a todo vetor de $W^\perp$.

Note que $v - proj_W(v)$ possui as mesmas propriedades de $proj_{W^\perp}(v)$, pois $v - proj_W(v) \in W^\perp$, pela propriedade 2 da Definição 3.1, e $v - (v - proj_W(v)) = proj_W(v) \in W$ que é ortogonal a todo vetor de $W^\perp$. Assim $v - proj_W(v) = proj_{W^\perp}(v)$.

A última informação do enunciado desse teorema é simplesmente o fato que $v - proj_{W^\perp}(v) = proj_W(v)$, junto com a equação

$$\langle v, v \rangle = \langle v - proj_W v, v - proj_W(v) \rangle + 2 \overbrace{\langle v - proj_W(v), proj_W(v) \rangle}^{=0} + \langle proj_W(v), proj_W(v) \rangle.$$

■

Observação 3.1

Na demonstração dessa proposição aprendemos uma coisa importante sobre a matriz de Gram de vetores linearmente independentes, ela tem determinante diferente de zero. Iremos usar essa informação no Teorema 4.1.

Agora provamos o resultado que havíamos prometido no início da seção: a $proj_W(v)$ é o vetor mais próximo de v dentro de W. Veja a seguinte proposição.

Proposição 3.2

Seja W um subespaço de dimensão finita de um espaço vetorial real V com produto interno e $v \in V$. Para todo $s \in W$ temos que

$$|v - proj_W(v)| \leq |v - s|.$$

Portanto $\inf_{s \in W} |v - s| = |v - proj_W(v)|$, ou seja, a distância entre v e W vale $|v - proj_W(v)|$.

Demonstração: Pela desigualdade de Cauchy-Schwartz temos que

$$|v - proj_W(v)||v - s| \geq |\langle v - proj_W(v), v - s \rangle|. \tag{4}$$

Mas

$$\langle v - proj_W(v), v - s \rangle = \langle v - proj_W(v), v \rangle = \langle v - proj_W(v), v - proj_W(v) \rangle, \tag{5}$$

pois $\langle v - proj_W(v), s \rangle = \langle v - proj_W(v), proj_W(v) \rangle = 0$, já que s e $proj_W(v)$ pertencem a W.

Assim pela desigualdade (4) e pela igualdade (5) temos

$$|v - proj_W(v)||v - s| \geq |\langle v - proj_W(v), v - proj_W(v)\rangle| = |v - proj_W(v)|^2.$$

Provando o resultado. ■

Na próxima seção obtemos a fórmula para distância de vetor a subespaço vetorial. Ela está totalmente conectada com a matriz de Gram, isso ocorre porque a matriz de Gram está intimamente conectada com a projeção, como vimos nessa seção.

4 Distância de vetor a subespaço

Finalmente estamos em condições de provar o resultado que dá a distância de um vetor v, que pertence a um espaço vetorial real com produto interno V, a um subespaço W de V com base $w_1, \ldots, w_n$.

Teorema 4.1

Seja V um espaço vetorial real com produto interno, $v \in V$ e W um subespaço de V com base $w_1, \ldots, w_n$. Se d é distância de v a W então

$$d^2 = |v - proj_W(v)|^2 = \frac{\det(Gram(w_1, \ldots, w_n, v))}{\det(Gram(w_1, \ldots, w_n))}.$$

Demonstração: Vimos na seção anterior que se $proj_W(v) = a_1 w_1 + \ldots a_n w_n$ então

$$\begin{pmatrix} \langle w_1, w_1\rangle & \ldots & \langle w_1, w_n\rangle \\ \vdots & \ddots & \vdots \\ \langle w_n, w_1\rangle & \ldots & \langle w_n, w_n\rangle \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \langle w_1, v\rangle \\ \vdots \\ \langle w_n, v\rangle \end{pmatrix}. \tag{6}$$

Agora vimos na Proposição 3.2 que $d = |v - proj_W(v)|$. Assim

$$d^2 = \langle v - proj_W(v), v - proj_W(v)\rangle = \langle v, v - proj_W(v)\rangle - \overbrace{\langle proj_W(v), v - proj_W(v)\rangle}^{=0}.$$

Portanto

$$d^2 = \langle v, v - a_1 w_1 - \ldots - a_n w_n\rangle = \langle v, v\rangle - a_1\langle v, w_1\rangle - \ldots - a_n\langle v, w_n\rangle.$$

Podemos acrescentar essa última equação ao sistema (6),

$$\begin{pmatrix} \langle w_1, w_1\rangle & \ldots & \langle w_1, w_n\rangle & 0 \\ \vdots & \ddots & \vdots & \vdots \\ \langle w_n, w_1\rangle & \ldots & \langle w_n, w_n\rangle & 0 \\ \langle v, w_1\rangle & \ldots & \langle v, w_n\rangle & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \\ d^2 \end{pmatrix} = \begin{pmatrix} \langle w_1, v\rangle \\ \vdots \\ \langle w_n, v\rangle \\ \langle v, v\rangle \end{pmatrix}.$$

Fazendo a expansão de Laplace na última coluna da matriz desse sistema para calcular seu determinante obtemos

$$\det \begin{pmatrix} \langle w_1, w_1\rangle & \ldots & \langle w_1, w_n\rangle & 0 \\ \vdots & \ddots & \vdots & \vdots \\ \langle w_n, w_1\rangle & \ldots & \langle w_n, w_n\rangle & 0 \\ \langle v, w_1\rangle & \ldots & \langle v, w_n\rangle & 1 \end{pmatrix} = 1. \det \begin{pmatrix} \langle w_1, w_1\rangle & \ldots & \langle w_1, w_n\rangle \\ \vdots & \ddots & \vdots \\ \langle w_n, w_1\rangle & \ldots & \langle w_n, w_n\rangle \end{pmatrix} \neq 0,$$

pois $w_1, \ldots, w_n$ formam uma base (Observação 3.1).

Assim esse último sistema tem solução única, $(a_1, \ldots, a_n, d^2)$, e o último coeficiente dessa solução única, d^2, pode ser calculado pela regra de Cramer como

$$\det\begin{pmatrix}\langle w_1,w_1\rangle & \dots & \langle w_1,w_n\rangle & \langle w_1,v\rangle\\ \vdots & \ddots & \vdots & \vdots\\ \langle w_n,w_1\rangle & \dots & \langle w_n,w_n\rangle & \langle w_n,v\rangle\\ \langle v,w_1\rangle & \dots & \langle v,w_n\rangle & \langle v,v\rangle\end{pmatrix}\Bigg/\det\begin{pmatrix}\langle w_1,w_1\rangle & \dots & \langle w_1,w_n\rangle\\ \vdots & \ddots & \vdots\\ \langle w_n,w_1\rangle & \dots & \langle w_n,w_n\rangle\end{pmatrix}=\frac{\det(Gram(w_1,\dots,w_n,v))}{\det(Gram(w_1,\dots,w_n))}.$$

■

Corolário 4.1

Seja V um espaço vetorial real com produto interno, a,b dois elementos de V e W_1,W_2 dois subespaços de V ambos de dimensão finita. Seja $w_1,\dots,w_n$ base do subespaço W_1+W_2. Se d é a distância entre os subespaços afins W_1+a e W_2+b então

$$d^2=\frac{\det(Gram(w_1,\dots,w_n,a-b))}{\det(Gram(w_1,\dots,w_n))}.$$

Em particular se $W_1=\{\vec{0}\}$ então $w_1,\dots,w_n$ é base de W_2 e a fórmula acima dá o quadrado da distância de a ao W_2+b.

Demonstração: Pela Observação 2.1, a distância de W_1+a a W_2+b é igual a distância de $a-b$ ao subespaço W_1+W_2. Agora o resultado que queremos provar segue do Teorema 4.1. ■

Se soubermos uma base para $W^\perp$ então podemos calcular a distância de um vetor v a W, sem calcular a base de W. Veja o seguinte corolário.

Corolário 4.2

Seja V um espaço vetorial real de dimensão finita com produto interno e $v_1,\dots,v_n$ uma base de $W^\perp$. Se d é distância de um vetor v ao subespaço vetorial W então

$$d^2=-\frac{\det(D(v_1,\dots,v_n,v))}{\det(Gram(v_1,\dots,v_n))},$$

onde $D(v_1,\dots,v_n,v)$ é uma matriz de ordem $n+1$ definida por

$$D(v_1,\dots,v_n,v)_{ij}=\begin{cases}Gram(v_1,\dots,v_n,v)_{ij}, & \text{para todo } (i,j)\neq(n+1,n+1)\\ 0, & \text{quando } (i,j)=(n+1,n+1)\end{cases}.$$

Demonstração: Como V tem dimensão finita então W e $W^\perp$ também têm dimensões finitas. Pela Proposição 3.2, temos que o quadrado da distância de v a $W^\perp$ vale $|v-proj_{W^\perp}(v)|^2$ e o quadrado da distância de v a W vale $d^2=|v-proj_W(v)|^2$. Pelo item 2 da Proposição 3.1 temos que $d^2=|v|^2-|v-proj_{W^\perp}(v)|^2$. Assim pelo Teorema 4.1 temos que

$$d^2=|v|^2-\frac{\det(Gram(v_1,\dots,v_n,v))}{\det(Gram(v_1,\dots,v_n))}. \tag{7}$$

Para concluir o corolário veja que podemos escrever

$$\det(Gram(v_1,\dots,v_n,v))=\det\begin{pmatrix}\langle v_1,v_1\rangle & \dots & \langle v_1,v_n\rangle & \langle v_1,v\rangle\\ \vdots & \vdots & \ddots & \vdots\\ \langle v_n,v_1\rangle & \dots & \langle v_n,v_n\rangle & \langle v_n,v\rangle\\ 0 & \dots & 0 & \langle v,v\rangle\end{pmatrix}+\det\begin{pmatrix}\langle v_1,v_1\rangle & \dots & \langle v_1,v_n\rangle & \langle v_1,v\rangle\\ \vdots & \vdots & \ddots & \vdots\\ \langle v_n,v_1\rangle & \dots & \langle v_n,v_n\rangle & \langle v_n,v\rangle\\ \langle v,v_1\rangle & \dots & \langle v,v_n\rangle & 0\end{pmatrix}=$$

$$\det(Gram(v_1,\ldots,v_n))|v|^2+\det(D(v_1,\ldots,v_n,v)).$$

Substituindo essa informação na equação (7) obtemos o resultado do enunciado. ■

5 Aplicações da fórmula do Teorema 4.1

Abaixo mencionamos algumas aplicações dessa fórmula obtida no Teorema 4.1 e das fórmulas dos seus corolários. Sabemos que algumas dessas aplicações são mais fáceis de serem obtidas apenas calculando a projeção de vetor em subespaço. Isso não quer dizer que essas fórmulas são ruins, apenas mostram que em casos bem particulares outros argumentos são melhores. Esse é o preço a se pagar por uma fórmula com variados usos. Existem outras situações onde essas fórmulas mostram seu real valor, como o Teorema de Müntz que descrevemos abaixo.

5.1 Distância entre ponto e o conjunto solução de um sistema no $\mathbb{R}^n$

Considere $y=(y_1,\ldots,y_n)$ um ponto no $\mathbb{R}^n$ e o subespaço afim que é solução do sistema

$$\begin{cases} a_{11}x_1+\ldots+a_{1n}x_n+d_1=0 \\ \vdots \\ a_{m1}x_1+\ldots+a_{mn}x_n+d_m=0 \end{cases}.$$

Suponha que $a_1=(a_{11},\ldots,a_{1n})$, $\ldots$, $a_m=(a_{m1},\ldots,a_{mn})$ sejam linearmente independentes e seja $b=(b_1,\ldots,b_n)$ uma das soluções desse sistema. Se $W+b$ é o subespaço afim que é solução desse sistema então $a_1,\ldots,a_m$ formam uma base de $W^\perp$. Como sabemos a base de $W^\perp$ então podemos usar a fórmula do Corolário 4.2 para calcular a distância d entre $y-b$ e W que, pela Observação 2.1, é a mesma distância entre y e $W+b$. Assim

$$d^2=-\frac{\det(D(a_1,\ldots,a_m,y-b))}{\det(Gram(a_1,\ldots,a_m))}.$$

Já chegamos na fórmula que queríamos. Agora observe o que ocorre quando só temos uma equação no sistema. Então

$$d^2=-\frac{\det(D(a_1,y-b))}{\det(Gram(a_1))}=-\frac{\det\begin{pmatrix}\langle a_1,a_1\rangle & \langle a_1,y-b\rangle \\ \langle y-b,a_1\rangle & 0\end{pmatrix}}{|a_1|^2}=\frac{\langle a_1,y-b\rangle^2}{|a_1|^2}.$$

Agora como b satisfaz $\langle a_1,b\rangle+d_1=0$ então

$$\langle a_1,y-b\rangle=a_{11}y_1+\ldots+a_{1n}y_n+d_1.$$

Assim

$$d=\frac{|\langle a_1,y-b\rangle|}{|a_1|}=\frac{|a_{11}y_1+\ldots+a_{1n}y_n+d_1|}{\sqrt{a_{11}^2+\ldots+a_{1n}^2}}.$$

Essa é a fórmula da distância de ponto a hiperplano no $\mathbb{R}^n$ que é o conjunto solução de um sistema com uma única equação. Note que daí tiramos as fórmulas de distância de ponto a reta no $\mathbb{R}^2$ e de ponto a plano no $\mathbb{R}^3$ descritas na introdução. É possível usar a fórmula do Corolário 4.1 diretamente para calcular a distância de ponto a hiperplano, ao invés do Corolário 4.2. Entretanto isso dá um pouco mais de trabalho. Veja abaixo como fica a demonstração da fórmula da distância de ponto a plano no $\mathbb{R}^3$ só com o Corolário 4.1.

5.2 Distância entre ponto e plano no $\mathbb{R}^3$ usando Corolário 4.1

Considere (x_0,y_0,z_0) um ponto no $\mathbb{R}^3$ e o plano $ax+by+cz+d=0$. Sabemos que (a,b,c) é o vetor normal a esse plano e qualquer vetor perpendicular a ele é um vetor diretor do plano. Escolhemos vetores v_1 e v_2 tais que $v_1\times v_2=(a,b,c)$, onde $v_1\times v_2$ é o produto vetorial de v_1 e v_2. Então eles não são múltiplos um do outro, pois seu produto vetorial não dá o vetor nulo e ambos são perpendiculares ao seu produto vetorial. Assim v_1,v_2 são vetores diretores do plano e também são linearmente independentes.

Note que o ponto

$$(x_1, y_1, z_1) = \left(-\frac{da}{a^2+b^2+c^2}, -\frac{db}{a^2+b^2+c^2}, -\frac{dc}{a^2+b^2+c^2}\right)$$

pertence ao plano $ax+by+cz+d=0$. Assim uma equação paramétrica desse plano é $(x_1, y_1, z_1) + \lambda v_1 + \mu v_2$.

Pelo Corolário 4.1, a distância de (x_0, y_0, z_0) ao plano $(x_1, y_1, z_1) + \lambda v_1 + \mu v_2$ vale

$$\sqrt{\frac{\det(Gram(v_1, v_2, v))}{\det(Gram(v_1, v_2))}}, \tag{8}$$

onde $v = (x_0 - x_1, y_0 - y_1, z_0 - z_1) = \left(x_0 + \frac{da}{a^2+b^2+c^2}, y_0 + \frac{db}{a^2+b^2+c^2}, z_0 + \frac{dc}{a^2+b^2+c^2}\right)$.

Agora, pela identidade de Lagrange,

$$\det(Gram(v_1, v_2)) = \langle v_1, v_1\rangle\langle v_2, v_2\rangle - \langle v_1, v_2\rangle^2 = |v_1 \times v_2|^2 = |(a,b,c)|^2 = a^2+b^2+c^2. \tag{9}$$

Note também que $Gram(v_1, v_2, v) = R^tR$, onde $R = (v_1, v_2, v)$ é uma matriz 3×3 cujas colunas são v_1, v_2 e v. Portanto

$$\det(Gram(v_1, v_2, v)) = \det(R^t)^2 = \det((v, v_1, v_2)^t)^2 = \langle v, v_1 \times v_2\rangle^2,$$

aqui estamos usando a fórmula do produto misto.

Como $v_1 \times v_2 = (a,b,c)$ e $v = \left(x_0 + \frac{da}{a^2+b^2+c^2}, y_0 + \frac{db}{a^2+b^2+c^2}, z_0 + \frac{dc}{a^2+b^2+c^2}\right)$ então

$$\langle v, v_1 \times v_2\rangle = \langle v, (a,b,c)\rangle = ax_0 + by_0 + cz_0 + d. \tag{10}$$

Portanto pelas equações (9) e (10) temos

$$\det(Gram(v_1, v_2, v)) = (ax_0+by_0+cz_0+d)^2 \quad \text{e} \quad \det(Gram(v_1, v_2)) = a^2+b^2+c^2.$$

Pela equação (8), temos a famosa fórmula de distância de ponto a plano no $\mathbb{R}^3$.

5.3 Distância entre planos no $\mathbb{R}^n$ (Caso particular do Corolário 4.1).

Considere as equações paramétricas dos planos $a + \lambda w_1 + \mu w_2$ e $b + \alpha w_3 + \beta w_4$ no $\mathbb{R}^n$. Qual é a distância entre eles ? Vimos, na Observação 2.1, que a distância entre $a + \lambda w_1 + \mu w_2$ e $b + \alpha w_3 + \beta w_4$ é a distância de $a - b$ ao subespaço W gerado por w_1, w_2, w_3, w_4. Podemos obter uma base de W a partir desse conjunto gerador e usar a fórmula do Corolário 4.1 para calcular a distância entre $v = a - b$ e W.

5.4 Volume do sólido gerado por vetores linearmente independentes.

Sejam $v_1, \ldots, v_k$ vetores linearmente independentes de um espaço vetorial real com produto interno. Considere o sólido $[v_1, \ldots, v_k] = \{x_1v_1 + \ldots + x_kv_k, \ \ x_i \in [0,1]\}$. Seu volume é definido indutivamente,

$$\text{vol}([v_1]) = |v_1| \quad \text{e} \quad \text{vol}([v_1, \ldots, v_i]) = \text{vol}([v_1, \ldots, v_{i-1}]) \times \text{dist}(v_i, W_{i-1}),$$

onde W_{i-1} é o espaço vetorial gerado por $v_1, \ldots, v_{i-1}$ e $\text{dist}(v_i, W_{i-1})$ é a distância de v_i ao W_{i-1}.

Por indução provamos que

$$\text{vol}([v_1, \ldots, v_i]) = \sqrt{\det(Gram(v_1, \ldots, v_i))}. \tag{11}$$

Para $i = 1$, temos que

$$\text{vol}([v_1]) = \sqrt{\det(Gram(v_1))} = \sqrt{\langle v_1, v_1\rangle} = |v_1|.$$

Suponha que (11) vale para $i = j$ e provemos para $i = j+1$. Por definição, $\text{vol}([v_1, \ldots, v_{j+1}]) = \text{vol}([v_1, \ldots, v_j]) \times \text{dist}(v_{j+1}, W_j)$. Pelo Teorema 4.1,

$$\text{dist}(v_{j+1}, W_j) = \frac{\sqrt{\det(Gram(v_1, \ldots, v_{j+1}))}}{\sqrt{\det(Gram(v_1, \ldots, v_j))}}.$$

Por hipótese de indução,

$$\text{vol}([v_1, \ldots, v_j]) = \sqrt{\det(Gram(v_1, \ldots, v_j))}.$$

Assim

$$\text{vol}([v_1, \ldots, v_{j+1}]) = \sqrt{\det(Gram(v_1, \ldots, v_j))} \times \frac{\sqrt{\det(Gram(v_1, \ldots, v_{j+1}))}}{\sqrt{\det(Gram(v_1, \ldots, v_j))}}.$$

Concluindo a indução e provando a equação (11).

5.5 O Teorema de Müntz

Uma das aplicações mais bonitas dessa fórmula do Teorema 4.1 é o Teorema de Müntz que descrevemos a seguir. Considere o espaço das funções reais contínuas com domínio $[0, 1]$, denotado por $C[0, 1]$, com o produto interno

$$\langle f(t), g(t)\rangle = \int_0^1 f(t)g(t)dt.$$

Considere as funções contínuas $t^{x_1}, t^{x_2}, t^{x_3}, \ldots$, onde $0 < x_1 < x_2 < x_3 < \ldots$. Podemos calcular a distância de t^k até o espaço gerado por $t^{x_1}, \ldots, t^{x_n}$ utilizando a fórmula do Teorema 4.1 com esse produto interno.

Obtém-se o seguinte resultado

$$\sqrt{\frac{\det(Gram(t^{x_1}, \ldots, t^{x_n}, t^k))}{\det(Gram(t^{x_1}, \ldots, t^{x_n}))}} = \frac{|k - x_1| \ldots |k - x_n|}{|k + x_1 + 1| \ldots |k + x_n + 1|(2k + 1)^{\frac{1}{2}}}.$$

É possivel provar que

$$\lim_{n\to\infty} \frac{|k - x_1| \ldots |k - x_n|}{|k + x_1 + 1| \ldots |k + x_n + 1|(2k + 1)^{\frac{1}{2}}} = 0,$$

para $k = 0, 1, 2, \ldots$, se e só se,

$$\sum_{i=1}^{\infty} \frac{1}{x_i} = \infty.$$

Isso mostra que as funções $1, t, t^2, \ldots,$ podem ser aproximadas por funções do espaço gerado por $t^{x_1}, t^{x_2}, \ldots$, se e somente se,

$$\sum_{i=1}^{\infty} \frac{1}{x_i} = \infty.$$

Consequentemente os polinômios também podem ser aproximados por funções do mesmo espaço. Mas sabemos pelo famoso teorema de Weiertrass que qualquer função contínua pode ser aproximada por um polinômio. Assim temos que qualquer função contínua pode ser aproximada por funções do espaço gerado por $t^{x_1}, t^{x_2}, \ldots$, se e somente se,

$$\sum_{i=1}^{\infty} \frac{1}{x_i} = \infty.$$

A aproximação entre funções aqui está ocorrendo com a norma do produto interno, a chamada norma 2. O Teorema de Müntz geral fala da aproximação com a norma do supremo. Veja seu enunciado na referência (CAROTHERS, 2009).

Referências

CAROTHERS, Neal Lamar (2009). **A Short Course on Approximation Theory**. Bowling Green State University.

WEXLER, Charles (1962). **Analytic Geometry: A Vector Approach**. Addison - Wesley Publishing Company.