Relatório do Projeto Final - Álgebra Linear Computacional

Carlos Eduardo Valladares da Mota

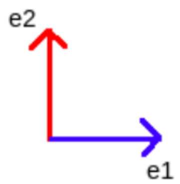
DRE: 119062426

1 Introdução

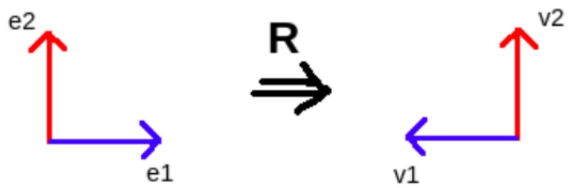
Antes de abordarmos a decomposição matricial na qual estamos interessados, vamos primeiramente motivá-la através de uma observação intuitiva sobre operadores lineares.

Vamos começar lembrando que, fixada uma base, todo operador linear possui uma única representação matricial, e vice-versa. Por simplicidade, vamos trabalhar com a base canônica, onde os vetores da base são colunas da matriz identidade correspondente ao nosso espaço de interesse. Também por simplicidade, vamos dar exemplos do $\mathbb{R}^2$ (plano), mas nossas afirmações se estendem para qualquer espaço R^n.

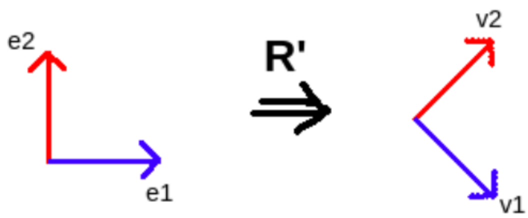
Inicialmente, então, temos nossos vetores da base sendo $e_1 = (1, 0)^T$ e $e_2 = (0, 1)^T$ como no desenho a seguir:



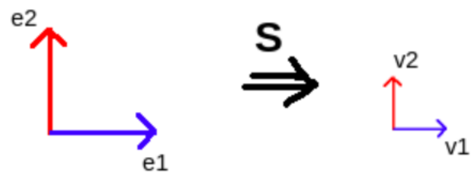
Vamos escolher novas posições para estes vetores. Digamos que queremos que e_1 esteja à esquerda de e_2, com mesmo tamanho. Podemos sair da base canônica e chegar nesta nova configuração dos vetores no espaço, aplicando uma matriz de reflexão R sob o eixo e_2:



Digamos que, em vez disso, quiséssemos rodar os eixos canônicos em 45°. Podemos fazer isso aplicando uma matriz de rotação R' que faz justamente isso:

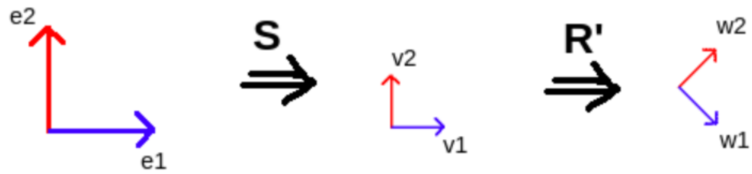


Mas e se quiséssemos mudar o comprimento dos vetores canônicos? Como, por exemplo, deixá-los com metade do tamanho? Podemos fazer isso através de uma matriz de escalonamento S:

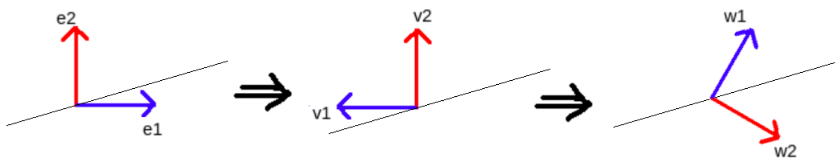


Estas são transformações lineares relativamente simples. Conhecemos algumas outras mais complicadas como: rotação com mudança de escala, reflexão sob uma reta, projeção, cisalhamento etc. Mas veja que interessante:

Para rotação com mudança de escala, podemos aplicar primeiro uma mudança de escala, e depois uma rotação:



Para a reflexão sob uma reta, podemos imaginar uma rotação seguida de uma reflexão:



Para a projeção, basta uma mudança de escala. Para o cisalhamento é bem mais difícil de imaginar este processo, mas não nos parece impossível. O que estas operações parecem sugerir é que: Dado um operador linear, eu consigo descrevê-lo usando uma mudança de escala e uma combinação de rotações/reflexões. Ou, na linguagem matricial: Dada uma matriz, é possível sempre decompô-la em uma matriz simétrica (mudanças de escala) e uma matriz ortogonal (rotações e reflexões). Para esta formulação matricial, vamos nos lembrar que o produto de matrizes ortogonais resulta em uma matriz ortogonal e, portanto, podemos juntar vários operadores de reflexão e rotação em um só.

Este resultado nos parece plausível, mas certamente não é óbvio/trivial. Vamos, portanto, tentar verificá-lo nas próximas seções. Começaremos definindo bem qual o problema que queremos resolver.

2 Formulação inicial do problema: a Decomposição Polar

Seja A uma matriz de dimensões $n \times n$, de coeficientes constantes, queremos encontrar uma decomposição de A em um produto de duas matrizes: à esquerda uma matriz Q **ortogonal**, de dimensões $n \times n$ e à direita uma matriz S **simétrica**, também de dimensões $n \times n$. Sucintamente, queremos encontrar a decomposição:

$$A_{n\times n} = Q_{n\times n}\, S_{n\times n} \tag{1}$$

Chamaremos esta decomposição da matriz A de **decomposição polar** de A. Posteriormente, faremos algumas alterações nesta formulação. Porém, por hora, dada a motivação de operadores lineares apresentada anteriormente, esta formulação nos parece adequada (e já generalizada do $\mathbb{R}^2$ para o $\mathbb{R}^n$).

3 Solução do problema via Teorema Espectral

Inicialmente conhecemos apenas A. Para tentarmos utilizar o teorema espectral para nos ajudar, precisamos garantir que a matriz a ser decomposta seja diagonalizável. Não temos esta garantia para A, então vamos analisar A^TA e, a partir da análise nesta matriz, ver se concluímos algo sobre Q ou S. Primeira afirmação: A^TA é uma matriz simétrica, e usaremos este fato mais adiante. Aqui uma rápida demonstração:

$$(A^TA)^T = A^T(A^T)^T = A^TA$$

Usando a equação (1), para A^TA:

$$\begin{aligned} A^TA &= (QS)^TQS \\ &= S^TQ^TQS \end{aligned}$$

Como Q é ortogonal, $Q^{-1} = Q^T$, e temos:

$$A^TA = S^TS$$

Mas S é simétrica, portanto $S^T = S$ e:

$$A^TA = S^TS = SS = S^2 \tag{2}$$

Sabemos, agora, como encontrar S^2 a partir de A. Estamos interessados em encontrar S a partir de S^2. Vamos utilizar um teorema da Álgebra Linear, cuja demonstração pode ser encontrada em diversas referências (como neste PDF): **Toda matriz simétrica é diagonalizável, e todos os seus autovalores são maiores ou iguais a zero**. S, por hipótese, é simétrica, e portanto podemos fazer a diagonalização:

$$S = Q'\Lambda'Q'^T$$

Sabemos que A^TA é simétrica, e que $S^2 = A^TA$, logo S^2 é simétrica e também podemos fazer sua diagonalização:

$$S^2 = Q''\Lambda''Q''^T$$

Queremos, então, que $S\ .\ S = S^2$. Substituindo a diagonalização das matrizes nesta igualdade, obtemos:

$$\begin{aligned} (Q'\Lambda'Q'^T)(Q'\Lambda'Q'^T) &= Q''\Lambda''Q''^T \\ Q'\Lambda'Q'^TQ'\Lambda'Q'^T &= Q''\Lambda''Q''^T \\ Q'\Lambda'\Lambda'Q'^T &= Q''\Lambda''Q''^T \end{aligned} \tag{3}$$

Para fazermos valer esta igualdade, igualamos $Q' = Q''$ e $\Lambda'\Lambda' = \Lambda''$. Lembrando que Λ' é diagonal, estamos apenas pedindo que $\lambda''_{ii} = \lambda'_{ii} * \lambda'_{ii}$ que são números reais. E podemos fazer isso de duas formas:

$$\lambda'_{ii} = \sqrt{\lambda''_i i}$$

ou

$$\lambda'_{ii} = -\sqrt{\lambda''_i i}$$

Para tentarmos garantir unicidade, vamos considerar apenas a primeira versão, com os valores não negativos. Deste jeito, padronizamos a forma de determinar S a partir de S^2:

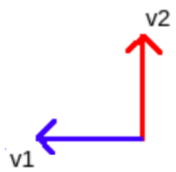
$$S = Q'' \sqrt{\Lambda''}\, Q''^T$$

Ou, em uma notação mais usual:

$$S = \sqrt{A^T A}$$

A imposição desta restrição, acarreta em uma mudança em nossa formulação inicial: a matriz S não é, mais, apenas simétrica, mas sim **simétrica positiva semi-definida**. Podemos afirmar isso devido ao seguinte teorema: uma matriz é positiva semi-definida se, e somente se, todos os seus autovalores são maiores ou iguais a 0.

Olhando de volta para o nosso contexto inicial de operadores lineares, estamos impedindo mudanças de escala com fatores negativos. Desta forma, ao observamos os vetores:



Saindo dos vetores canônicos, saberemos que v_1 e v_2 foram alcançados através de uma reflexão sob o eixo vertical, e não uma mudança de escala horizontal com fator negativo.

Sabemos, portanto, determinar S a partir de A. Vamos, agora, tentar determinar Q a partir de A e S. Como queremos que $A = QS$, podemos tentar isolar Q nesta expressão, e assim expressamos a matriz em termos de objetos já conhecidos. Pela equação (1), se **S for inversível** então:

$$\begin{aligned} A &= QS \\ AS^{-1} &= QSS^{-1} \\ AS^{-1} &= Q \end{aligned} \tag{4}$$

E encontramos uma forma de determinar Q. Pergunta: Q é realmente ortogonal? A resposta é que sim. Veremos logo a frente que não estamos tão interessados nesta forma de construir Q, e por isso não investiremos, aqui, muito tempo neste resultado. Mas, caso queira ver o porquê de Q ser ortogonal, há uma seção extra dedicada a esta demonstração ao final deste documento.

Encontramos, então, uma forma de encontrar a decomposição polar de uma matriz A: Fazemos

$$\begin{aligned} Q &= AS^{-1} \\ S &= \sqrt{A^T A} \end{aligned}$$

Mas por que não estamos muito interessados neste resultado? Porque ele depende que S seja inversível, e esta imposição é muito forte. Pedir que S seja inversível faz com que, por exemplo, não tenhamos mais os operadores de projeção. Além disso, nos apoiamos no fato de que S é diagonalizável para fazermos sua construção, mas e se quisermos estender nossa definição para um operador A de dimensões $n \times m$? Nesta situação, S será retangular e não mais diagonalizável.

Para resolvermos estas limitações e conseguirmos não apenas resolver nosso problema inicialmente proposto, mas também generalizarmos nossa solução para um problema mais genérico, vamos utilizar uma estratégia diferente.

4 Solução do problema via SVD

Vamos analisar novamente nossa formulação do problema, porém agora sob outro olhar. Para construirmos S na seção anterior, utilizamos a diagonalização de $A^T A$. Por que fizemos isso para $A^T A$, mas não para A? Se fizéssemos para A, poderíamos escrever a solução diretamente em termos da nossa matriz original, em vez de passarmos por esta matriz intermediária. A resposta está tanto nas características de A quanto no método de decomposição: A é uma matriz **qualquer**, e para fazermos a decomposição espectral, precisávamos que ela fosse **diagonalizável**. Não tínhamos esta garantia, e portanto precisávamos de $A^T A$.

Agora, porém, não precisamos de qualquer propriedade de A para utilizarmos o SVD! Desta forma, podemos determinar $A = U\Sigma V^T$ e vemos que:

$$
\begin{aligned}
S &= \sqrt{A^T A} \\
&= \sqrt{(V\Sigma U^T)(U\Sigma V^T)} \\
&= \sqrt{V\Sigma\Sigma V^T} \\
&= \sqrt{V\Sigma^2 V^T} \\
&= V\Sigma V^T
\end{aligned}
$$

Expressando S diretamente em termos de A, o que traz a economia computacional de não precisarmos calcular $A^T A$. Note que, na operação de extração de raiz, mantivemos a mesma restrição de sinais da seção anterior. Ou seja, Σ possui apenas valores não negativos. Desta forma, S permanece sendo positiva semi-definida.

Agora, para determinarmos Q, podemos repetir nossa estratégia anterior de isolar Q na equação $A = QS$. Desta vez, porém, temos a vantagem de que A e S podem ser descritos relativos a uma mesma decomposição (a decomposição SVD de A):

$$
\begin{aligned}
A &= QS \\
U\Sigma V^T &= QV\Sigma V^T
\end{aligned}
$$

Multiplicando por V à direita e usando o fato de que V é ortogonal:

$$
\begin{aligned}
U\Sigma V^T V &= QV\Sigma V^T V \\
U\Sigma &= QV\Sigma
\end{aligned}
$$

E queremos um Q que satisfaça esta equação. Note que não podemos continuar a passagem multiplicando pela inversa de Σ pelos dois lados, já que Σ pode não ser inversível (podemos ter valores singulares iguais a 0). Analisando a última equação, podemos ver que $Q = UV^T$ parece funcionar enquanto solução:

$$
\begin{aligned}
U\Sigma &= QV\Sigma \\
U\Sigma &= (UV^T)V\Sigma \\
U\Sigma &= U\Sigma
\end{aligned}
$$

E determinamos uma expressão para Q sem precisarmos impor qualquer restrição adicional em nosso problema. Podemos nos fazer a mesma pergunta de novo: Q é realmente ortogonal? Sim, e desta vez podemos utilizar um teorema para provarmos este teorema em uma linha: O produto de 2 matrizes ortogonais resulta em uma matriz ortogonal. U é ortogonal, V^T também é. Logo, o produto destas matrizes, que é Q, também é ortogonal.

Adicionalmente, desta vez, usando o fato de que a decomposição SVD funciona para matrizes retangulares, resolvemos também o problema para A retangular, não apenas quadrada. Ou seja, resolvemos a forma mais genérica da formulação inicial do nosso problema:

> Sendo A uma matriz de dimensões $n \times m$, de coeficientes constantes, podemos encontrar uma decomposição de A em um produto de duas matrizes: à esquerda uma matriz Q **ortogonal**, de dimensões $n \times m$ e à direita uma matriz S **simétrica positiva semi-definida** de dimensões $m \times m$. Sucintamente:

$$A_{n \times m} = Q_{n \times m}\ S_{m \times m} \tag{5}$$

E encontramos uma forma de fazer esta outra decomposição polar. Sabendo $A = U\Sigma V^T$, tomamos:

$$
\begin{aligned}
Q &= UV^T \\
S &= V\Sigma V^T
\end{aligned}
$$

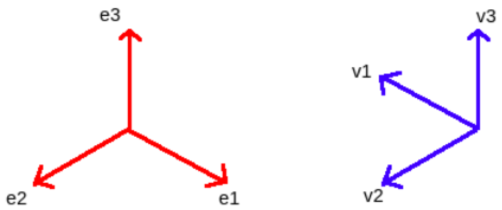
Não perdemos nenhum operador no caminho, e ainda englobamos operadores que atuam entre diferentes espaços. Caso não esteja claro que as dimensões dos produtos batem (agora que estamos operando com matrizes retangulares), aqui estão as operações com as dimensões explícitas:

$$
\begin{aligned}
A_{n \times m} &= U_{n \times k}\ \Sigma_{k \times k}\ V^T_{k \times m} \\
Q_{n \times m} &= U_{n \times k}\ V^T_{k \times m} \\
S_{m \times m} &= V_{m \times k}\ \Sigma_{k \times k}\ V^T_{k \times m}
\end{aligned}
$$

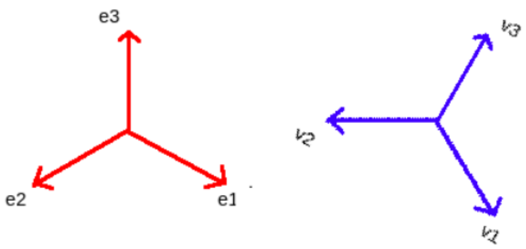
5 Exemplo de aplicação: posicionamento geográfico

Vamos imaginar que estamos em uma base aérea, e iremos decolar com um helicóptero. Este helicóptero irá percorrer uma região ao redor da base área, dentro de um raio de 6km. Como isso é 0,1% do raio da terra, podemos considerar que nosso espaço de navegação é plano (mas sempre bom lembrar que a terra não é plana).

Se colocarmos um sistema de coordenadas de referência na base área (representado em vermelho), e outro no helicóptero (representado em azul), sabemos que para sair de um sistema para o outro nos bastam transformações ortogonais (pois não estamos dilatando o espaço-tempo, logo só pode ter acontecido troca de orientação), como por exemplo:



E em um outro momento:



Vamos, então, simular uma navegação. Começaremos com ambos o helicóptero e a base área orientados da mesma forma, que será alinhados com os eixos canônicos do R^3. Sendo B a matriz de orientação da base aérea e H a matriz de orientação do helicóptero, temos inicialmente:

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} ; H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

O helicóptero, agora, sairá da base e começará uma ronda. Para simplificar nossa análise, vamos considerar que ele está andando apenas em círculos, no sentido anti-horário. Nosso giroscópio eletrônico, dentro do helicóptero, será responsável por indicar em que direção estamos relativamente a base aérea a qualquer momento, e é atualizado de tempos em tempos. Para fazer isso, ele determinará quais rotações a aeronave está fazendo, e aplicará a operação inversa (a transposta da matriz) para preservar a orientação original. Ou seja, se fizemos uma rotação de 45°:

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \rightarrow (R_{45^o}\ H) = \begin{bmatrix} cos(45^o) & -sen(45^o) & 0 \\ sen(45^o) & cos(45^o) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$(R_{45^o}\ H) = \begin{bmatrix} cos(45^o) & -sen(45^o) & 0 \\ sen(45^o) & cos(45^o) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

aplicando uma rotação de -45° voltamos a orientação original:

$$B = (R_{45^o}^{-1}\ H) = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Ou seja, recuperamos a orientação original relativa à base. Porém, esta matriz de rotação R, que estaremos calculando a todo momento, depende de sensores. Ou seja, está suscetível a erro, e portanto não é perfeitamente calculada. Estamos no nosso exemplo, então, calculando uma R que na verdade se assemelha a uma matriz do tipo:

$$\begin{bmatrix} cos(\theta) + \epsilon_1 & -sen(\theta) + \epsilon_2 & \epsilon_3 \\ sen(\theta) + \epsilon_4 & cos(\theta) + \epsilon_5 & \epsilon_6 \\ \epsilon_7 & \epsilon_8 & 1 + \epsilon_9 \end{bmatrix}$$

Onde cada ϵ é um erro de medição, por exemplo. Estes erros costumam ser pequenos, mas existem. O quanto eles podem nos atrapalhar? Vamos fazer o experimento:

5.1 Ronda no Helicóptero - usando os dados originais dos sensores

Vamos instanciar alguns números para nosso problema, para percebermos melhor (em escala e contexto) o que está acontecendo. Digamos que o nosso helicóptero está voando em um trajeto circular, onde o centro é a base área. A distância dele até a base aérea é de 6km, e ele está viajando a uma velocidade de aproximadamente 150km/h. Fazendo algumas contas, podemos determinar que a velocidade angular do helicóptero é de 0,00695 rad/s, ou seja, ele demora aproximadamente 15 minutos para completar uma volta completa (supondo que a cada instante de tempo a mesma distância é percorrida, sempre).

Vamos considerar que nosso giroscópio atualiza a uma taxa de 20 vezes por segundo. Logo, ele irá medir que o helicóptero está fazendo uma rotação de $\frac{0{,}00695}{20} = 0.0003475\ rad$ a cada instante de tempo. Logo, iremos aplicar sempre a mesma matriz de rotação:

$$\begin{bmatrix} cos(0.0003475) & -sen(0.0003475) & 0 \\ sen(0.0003475) & cos(0.0003475) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Se estivéssemos medindo perfeitamente! Mas não estamos, logo, há um erro nos ângulos medidos. Vamos supor que o erro é o mesmo em cada instante de tempo $\epsilon_1 = \epsilon_2 = ... = \epsilon_9$, e que este erro é distribuído de forma normal (com média 0 e variância 1). Para nós, 1 radiano impacta na escala de 6km. Vamos ser razoáveis, então, e dizer que os sensores erram na ordem dos **décimos de milímetros**, e portanto multiplicaremos este erro por 10^{-7} (0,1mm = 10^{-7}km). A cada iteração, portanto, teremos uma matriz:

$$\begin{bmatrix} cos(0.0003475) + 10^{-7}\epsilon & -sen(0.0003475) + 10^{-7}\epsilon & 10^{-7}\epsilon \\ sen(0.0003475) + 10^{-7}\epsilon & cos(0.0003475) + 10^{-7}\epsilon & 10^{-7}\epsilon \\ 10^{-7}\epsilon & 10^{-7}\epsilon & 10^{-7}\epsilon \end{bmatrix}$$

A cada instante de tempo t (que passa a cada 1/20 segundo) vamos fazer $H_t = R_t * H_{t-1}$ (posição atual é a posição anterior transformada pela rotação deste instante de tempo), e idealmente podemos determinar os vetores da base aérea B' fazendo $B' = H_t^T * H_t$ (pois a inversa da ortogonal é a transposta). Vamos ver o que ocorre com a posição da base área de acordo com nosso giroscópio ($B' = H_t^T * H_t$) ao longo do tempo:

Inicialmente:

$$B' = B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Após 1 segundo:

$$B' = \begin{bmatrix} 1.0 & 7.96109e-7 & 8.07615e-7 \\ 7.96109e-7 & 1.0 & 7.8485e-7 \\ 8.07615e-7 & 7.8485e-7 & 1.0 \end{bmatrix}$$

estamos errando na ordem de décimo de milímetros, tudo bem (afinal, somos um helicóptero)

Após 1 minuto:

$$B' = \begin{bmatrix} 0.999967 & -6.11266e-6 & -2.23233e-5 \\ -6.11266e-6 & 1.0 & -2.01337e-6 \\ -2.23233e-5 & -2.01337e-6 & 0.999986 \end{bmatrix}$$

estamos errando na ordem dos milímetros, tudo sob controle (por enquanto)

Após 1 hora:

$$B' = \begin{bmatrix} 0.9996 & -0.000151886 & -0.000513424 \\ -0.000151886 & 0.999407 & -0.000413023 \\ -0.000513424 & -0.000413023 & 0.999503 \end{bmatrix}$$

estamos errando na ordem dos metros, é contornável já que as estruturas de heliportos têm dezenas ou centenas de metros. Mas já é preocupante.

Após 1 dia:

$$B' = \begin{bmatrix} 1.00045 & -0.000297347 & 0.00118061 \\ -0.000297347 & 0.997759 & -0.00539464 \\ 0.00118061 & -0.00539464 & 0.99911 \end{bmatrix}$$

estamos errando nas dezenas de metros, já é perigoso para momentos críticos (como combustível no limite).

Após 1 mês:

$$B' = \begin{bmatrix} 0.992074 & 0.0177587 & 0.0215562 \\ 0.0177587 & 0.981961 & 0.0147345 \\ 0.0215562 & 0.0147345 & 0.987021 \end{bmatrix}$$

estamos errando centenas de metros, e você pode imaginar o que continuará ocorrendo daqui em diante.

A cada momento, nossa matriz R **quase** ortogonal é uma boa aproximação (como vimos nos primeiros instantes de tempo). Porém, o acumulado destas operações carrega o erro de todas elas, e depois de um certo tempo já temos erro o suficiente para o helicóptero achar que está navegando paralelo ao chão, quando na verdade está indo de encontro a ele. Como podemos mitigar o problema? As 6 primeiras páginas deste relatório indicam uma possível solução.

5.2 Ronda no Helicóptero - transformando os dados originais dos sensores

Gostaríamos que nossa R **quase** ortogonal fosse, de fato, ortogonal (até onde o erro de float permitir). Podemos usar a decomposição polar da matriz para nos ajudar! Sabemos que podemos decompor R em QS, onde Q é ortogonal e S é simétrica e representa mudança de escala. Mas sabemos que, idealmente, R é puramente ortogonal, e o termo S da decomposição é apenas erro de medição.

Podemos, então, adotar a seguinte estratégia: Cada vez que medirmos R, calculamos a sua decomposição polar $R = QS$, descartamos o termo S, e usamos $R = Q$. Vamos ver qual o impacto exercido no nosso mesmo problema anterior:

Inicialmente:

$$B' = B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Após 1 segundo:

$$B' = \begin{bmatrix} 1.0 & 4.15064e-16 & 3.2694e-17 \\ 4.15064e-16 & 1.0 & 6.92328e-16 \\ 3.2694e-17 & 6.92328e-16 & 1.0 \end{bmatrix}$$

Após 1 minuto:

$$B' = \begin{bmatrix} 1.0 & 4.21528e-15 & 8.1191e-15 \\ 4.21528e-15 & 1.0 & 2.87544e-15 \\ 8.1191e-15 & 2.87544e-15 & 1.0 \end{bmatrix}$$

Após 1 hora:

$$B' = \begin{bmatrix} 1.0 & -5.59074e-14 & -1.5461e-14 \\ -5.59074e-14 & 1.0 & -3.38434e-14 \\ -1.5461e-14 & -3.38434e-14 & 1.0 \end{bmatrix}$$

Após 1 dia:

$$B' = \begin{bmatrix} 1.0 & -6.0573e-13 & 4.24953e-13 \\ -6.0573e-13 & 1.0 & 6.02483e-14 \\ 4.24953e-13 & 6.02483e-14 & 1.0 \end{bmatrix}$$

E não estamos errando nem na escala de **milésimos de *nanômetro***! Certamente calcular o SVD de R a todo momento não é um dos algoritmos mais espertos, e simular tempos maiores que 1 dia começa a levar algum tempo computacional. Mas quando se tem o algoritmo rodando em tempo real, é relativamente plausível (ainda que existam alternativas mais eficientes). Porém, inegavelmente, observamos que uma eventual manutenção no giroscópio só precisaria ser feita depois de muito mais tempo de uso do que no caso inicial, onde usávamos os dados originais de R.

6 Conclusão

Vimos que, além de intuitiva quando analisada sob a ótica de transformações lineares, a decomposição polar é mais genérica do que poderíamos pensar inicialmente, e possui aplicações práticas bem interessantes. Existem outros problemas relacionados com esta decomposição, como o problema de Procustes Ortogonal, a outros contextos úteis de aplicação desta decomposição. Não entramos no tópico de algoritmos para calcular esta decomposição (além dos imediatos, derivados diretamente das demonstrações), mas existem muitas propostas com diferentes aplicações (no nosso exemplo do helicóptero, não precisaríamos calcular S, apenas Q).

Bônus: Caso você tenha ficado curioso com a decomposição polar de um cisalhamento (que dizemos ser estranha, mas não impossível), aqui está o resultado numérico (pois já adianto que a expressão simbólica é um pouco monstruosa), caso queira passar o seu tempo livre pensando no que acontece nessa transformação:

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.894427 & 0.447214 \\ -0.447214 & 0.894427 \end{bmatrix} \begin{bmatrix} 0.894427 & 0.447214 \\ 0.447214 & 1.34164 \end{bmatrix}$$

Apêndice

1) Demonstração que Q, via equação (4), é ortogonal

Se Q é ortogonal, então $Q^TQ = QQ^T = I$. Usando a equação (4):

$$\begin{aligned} Q^TQ &= (AS^{-1})^T AS^{-1} \\ &= {S^{-1}}^T A^T AS^{-1} \end{aligned}$$

Utilizamos mais um teorema: Se $S = S^T$, então $S^{-1} = {S^{-1}}^T$. Disto, segue:

$$Q^TQ = S^{-1}A^TAS^{-1}$$

Mas sabemos que $A^TA = S^2 = SS$:

$$\begin{aligned} Q^TQ &= (S^{-1}S)(SS^{-1}) \\ &= II \\ &= I \end{aligned}$$

Usando novamente as propriedades da transposta e da inversa (em ambas as direções), podemos mostrar que $QQ^T = I$:

$$\begin{aligned} QQ^T &= AS^{-1}(AS^{-1})^T \\ &= AS^{-1}{S^{-1}}^T A^T \\ &= AS^{-1}S^{-1}A^T \\ &= A(SS)^{-1}A^T \\ &= A(A^TA)^{-1}A^T \\ &= AA^{-1}{A^T}^{-1}A^T \\ &= II \\ &= I \end{aligned}$$

Da terceira para a quarta linha, usamos a propriedade da distribuição da inversa sob 2 matrizes (5ª propriedade nesta lista), mas no "sentido contrário". Assim, provamos que Q é ortogonal.

Pudemos utilizar a inversa de A na antepenúltima linha desta conta pois ao afirmarmos que S é inversível, garantimos que A também é inversível. Aqui uma rápida demonstração:

Afirmamos que S é inversível, logo SS é inversível (usando a mesma propriedade da distribuição da inversa, referenciada 2 parágrafos acima). Mas $A^TA = SS$, logo A^TA é inversível. Usando a propriedade da inversa $(A^TA)^{-1} = A^{-1}{A^T}^{-1}$, e temos garantida a existência da inversa de A e de sua transposta.