Álgebra Linear Numérica

Notas de Aula

Ricardo Fabbri

23 de Outubro de 2018

Aula 1

O que é uma matriz?

Objetivos

- Aula introdutória informal dando visão geral e uma contextualização do curso
- Motivar alunos de engenharia da computação da importância da disciplina ao final do curso, no contexto de modelagem computacional do IPRJ/UERJ.
- O que é álgebra linear numérica: explicar termos "álgebra", "linear", e "numérica". Por quê são importantes.
- Matrizes como representação numérica de transformações lineares
- Critério de avaliação

O que é uma matriz?

- Uma tabela retangular de números, porém não só.
- Associados à tabela há uma "álgebra": operações algébricas usuais de multiplicação matriz-vetor e matriz-matriz vistas no ensino médio, que parecem arbitrárias.
- Quando falamos de matrizes em engenharia e ciência, sempre incluimos, portanto, tais operações.

- O curso se chama Álgebra Linear Numérica. Até agora, então, falamos algo de "numérico" (matrizes são um monte de números) e de "algébra" (operações). E o "linear"?
- Matrizes são representações numéricas de transformações lineares.

O que é "Linear"?

- Ao ouvir a palavra "linear", a maioria das pessoas tem em mente uma reta ou um plano.
- "Linearidade" em matemática é uma forma de "simplicidade".
- Esta simplicidade se manifesta em diversos níveis: o geométrico, o simbólico, e o numérico:

Simbólico: No plano simbólico ou algébrico, explica-se os fenômenos de forma operacional e abstrata, sem muita intuição conceitual sobre o significado dos símbolos ou operações.

Definição 1. (Simbolicamente Linear) *Muitos matemáticos irão definir "linear" como a qualidade de uma função ou transformação L tal que*

$$\begin{cases} L(u+v) = L(u) + L(v) & (1.1) \\ L(\alpha v) = \alpha L(v), & (1.2) \end{cases}$$

para todos vetores u e v, e escalares α.

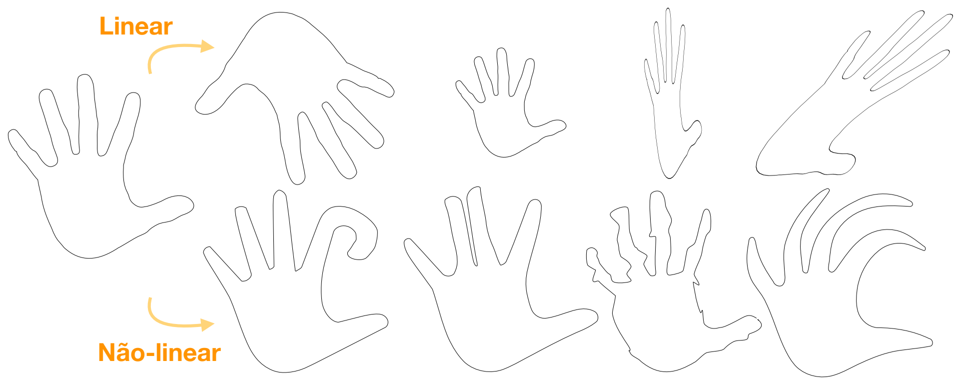
- Tal definição, apesar de sucinta e formalmente elegante, tem pouca concretude imediata.
- "Linear" aqui significa apenas uma simplicidade formal, ou seja, as manipulações algébricas são simples.

Geométrico: No plano que aqui chamamos de "geométrico", "conceitual", ou "visual", os fenômenos são explicados de maneira concreta, mais próxima do objetivo ou de aplicações reais.

No mundo real, o que observamos a princípio não é numérico nem simbólico: uma bola girando, um fluido em movimento, ou um feixe de nêutrons. Muitas vezes, conceitos próximos da realidade, independentes das convenções de símbolos ou números, são difíceis de formalizar e, portanto, não são enfatizados na maioria dos cursos de álgebra linear.

Definição 2. (Geometricamente Linear) *Chama-se "Linear" qualquer operação geométrica que combina quaisquer das seguintes operações simples: girar, esticar e refletir.*

- Tais operações são globais e realizadas sobre um mesmo ponto central (origem)
- O esticamento pode ser realizado por fatores diferentes ao longo de eixos arbitrários, e inclui achatamentos
- Geralmente, esta definição é um teorema na teoria da álgebra linear, visto apenas no final de um primeiro curso. Trata-se do teorema da SVD – *Singular Value Decompositon*, ou decomposição em valores singulares. Trata-se de um análogo de autovalores e autovetores, e é o teorema mais importante a ser explorado neste curso.
- Como exemplo, temos a seguinte figura:



A mão à esquerda sofre transformações lineares na linha acima, e não-lineares na linha abaixo. As transformações acima são lineares pois consistem em girar ou esticar cada ponto da mão globalmente pelo mesmo ângulo ou fatores de esticamento. Já as transformações na linha abaixo são não-lineares pois os pontos ou sofrem deformações locais (diferentes para cada ponto da mão), ou não são meramente girados e esticados.

Numérico: Já afirmamos no início da aula que uma matriz é a representação numérica de uma transformação linear, seja esta pensada pela definição geométrica ou simbólica.

- Para maior compreensão do significado dos números e operações de uma matriz, na próxima aula revisaremos o conceito de *coordenadas*.
- Também revisaremos como encontrar os números em uma matriz para uma dada transformação linear

O teorema da SVD pode ser descrito numericamente em termos de matrizes:

Teorema 1. *(SVD numérico - informal) Toda matriz A de tamanho $m \times n$ pode ser escrita na forma:*

$$\overbrace{A}^{\text{Qualquer matriz}} = \underbrace{V}_{\text{Ortogonal}}\underbrace{\Lambda}_{\text{Diagonal}}\underbrace{U^\top}_{\text{Ortogonal}}$$

Ou seja, toda matriz é uma matriz diagonal junto com matrizes ortogonais. A matriz diagonal realiza esticamento em cada eixo, e as matrizes ortogonais realizam rotações e reflexões.

- O tamanho das matrizes é

$$A_{m\times n} = V_{m\times m}\,\Lambda_{m\times n}\,U_{n\times n}^\top \tag{1.3}$$

- Portanto toda matriz é, de certa forma, equivalente a uma matriz diagonal de mesmo tamanho.
- Matrizes diagonais são matrizes da seguinte forma (exemplo 3×3):

$$\Lambda = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}, \tag{1.4}$$

as quais transformam pontos $p = (x, y, z)$ da seguinte forma:

$$\Lambda p = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} ax \\ by \\ cz \end{bmatrix}, \tag{1.5}$$

ou seja, esticam cada ponto ao longo de eixos por fatores a, b e c.

- Cada fator de esticamento a, b ou c pode ser zero, o que faz com que a matriz achate completamente as direções (dimensões) correspondentes.

Observações

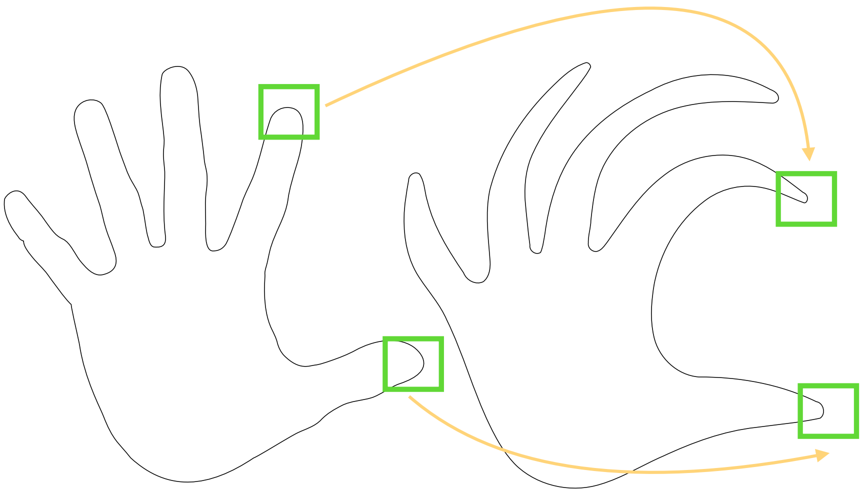
- A translação é não-linear. Pode-se mostrar que, mesmo sendo uma operação geométrica simples, a translação não tem as propriedades de simplicidade algébrica como na definiçao simbólica acima, e também não pode ser expressa numericamente como uma multiplicação de matrizes.
- Com alguns artifícios, é possível realizar translação com multiplicação de matrizes, porém é preciso operações não-lineares (projeção).

Para quê servem as transformações lineares, se elas são tão limitadas?

- O mundo real é não-linear
- O "Linear" foi inventado como uma simplificação do não-linear, mas em si não tem aplicação prática.
- O linear não faz sentido sem o não-linear.
- O linear só existe como artefato para modelar o não-linear.
- O paradigma é tomar um fenômeno real (e não-linear) e linearizá-lo, para usar as ferramentas deste curso.
- Em Cálculo 1, estudamos funções não-lineares de uma variável usando a derivada como linearização local (reta tangente). A reta tangente é uma transformação linear de uma variável.
- Daí a importância de se estudar cursos como análise no $\mathbb{R}^n$, onde fenômenos não-lineares são modelados tendo como ferramenta básica a álgebra linear.

- Álgebra linear também é o primeiro contato do aluno com n dimensões, para $n > 3$, e tem importância direta na prática, por exemplo em inteligência artificial, *machine learning*.

- Por exemplo, o paradigma de cálculo e análise no $\mathbb{R}^n$ é estudar o não-linear realizando linearizações locais (diferenciação). Dessa forma, para se deformar a mão de forma mais geral, como na figura abaixo, pode-se usar várias transformações lineares com fatores de esticamento e rotação diferentes para cada ponto:



- Em cada quadrado verde, temos uma deformação próxima de uma rotação e um esticamento.

- Curiosidade: a deformação não-linear da mão é uma função de $\mathbb{R}^2 \to \mathbb{R}^2$ estudada em Análise no $\mathbb{R}^n$. A aproximação da deformação de cada quadrado verde é chamada de diferencial da função não-linear, e sua forma numérica (matricial) é estudada em álgebra linear numérica.

Tarefas

- Site: http://wiki.nosdigitais.teia.org.br/ALN

- Realizar tarefa 0 e tarefa 1 (datas no site).

Exercícios

1. Explique quando uma função $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x)$, *i.e.*, $y = f(x)$, seria considerada linear, tanto em termos simbólicos quanto geométricos. Qual seria a fórmula geral dessa função?

Aula 2

Revisão de Álgebra Linear

Objetivos

- Relebrar representação numérica de vetores
- Revisar álgebra linear não-numérica ("simbólica")
- Estilo: revisão aprofundada para alunos que ja viram a disciplina. Notação solta adequada a uma revisão.

Observações Iniciais

- A álgebra liner numérica começa com o conceito de escalares e coordenadas.
- Para haver coordenadas, é necessária a escolha de uma base
- O conceito de espaço vetorial (simbólico) permite isolar as operações que não dependem da escolha de uma base, daquelas que dependem.

Definição 3. (Espaço Vetorial Simbólico) *Um espaço vetorial $\mathcal{V}$ sobre um corpo K é uma quádrupla $\mathcal{V} = (V, K, +, \cdot)$ de elementos:*

$$\mathcal{V} \quad \begin{cases} V\text{: conjunto cujos elementos são chamados "vetores"} \\ K\text{: conjunto cujos elementos são chamados "escalares"} \\ +\text{: soma de vetores} \\ \cdot\text{: multiplicação vetor-escalar,} \end{cases}$$

onde "+" satisfaz:

$$
\begin{aligned}
+ : V \times V &\to V && (2.1)\\
(v, w) &\mapsto v + w, && (2.2)
\end{aligned}
$$

$$
\begin{aligned}
& v + w = w + v && (2.3)\\
& v + (u + w) = (v + u) + w && (2.4)\\
& \exists!\ 0 \in V \ \textit{tal que}\ 0 + v = v,\ \forall v \in V && (2.5)\\
& \forall v \in V, \exists!\ w \in V \ \textit{tal que}\ v + w = 0, \textit{simbolizado por}\ -v && (2.6)
\end{aligned}
$$

e onde "·" satisfaz:

$$
\begin{aligned}
\cdot : K \times V &\to V && (2.7)\\
(\alpha, w) &\mapsto \alpha \cdot w, && (2.8)
\end{aligned}
$$

$$
\begin{aligned}
& 0v = 0 && \textit{o escalar 0 vezes um vetor é o vetor nulo} && (2.9)\\
& 1v = v && \textit{o escalar 1 vezes um vetor é o próprio vetor} && (2.10)\\
& (\alpha\beta)v = \alpha(\beta v) && \textit{multiplicação escalar} \cdot \textit{escalar é associativa com a escalar} \cdot \textit{vetor} && (2.11)\\
& \alpha(v + w) = \alpha v + \alpha w && \textit{escalar} \cdot \textit{vetor é distributiva com vetor} + \textit{vetor} && (2.12)\\
& (\alpha + \beta)v = \alpha v + \beta v && \textit{escalar} + \textit{escalar é distributiva com escalar} \cdot \textit{vetor.} && (2.13)
\end{aligned}
$$

e K satisfaz as propriedades de uma estrutura algébridca chamada "corpo", que abstrai as operações de multiplicação, soma, subtração e divisão possíveis com números, geralmente $\mathbb{R}$ ou $\mathbb{C}$.

- Aos programadores em C++: a definição de espaço vetorial simbólico é como uma *classe*, que define as partes e as operações permitidas.
- Na definição simbólica um vetor é apenas caracterizado pelas operações que se pode realizar com ele, e as propriedades de tais operações.
- Até o momento, a única coisa "numérica" nesta definição é o corpo de escalares K. Os vetores ainda são entidades simbólicas que podem ser somadas e multiplicadas por escalar.
- A teoria simbólica parece artificialmente trivial – todos já sabemos as propriedades convencionais de soma de vetores e multiplicação por escalar.

- No entanto, a teoria simbólica está mais próxima do conceitual que a teoria numérica, pois isola os fatores que não dependem da escolha de um sistema de coordenadas.
- A conexão com o mundo numérico consiste em associar números a vetores. Para tanto, é importante definir base e, para isso, independência linear.
- Iremos assumir que V e W são espaços vetoriais.

Definição 4. *Uma combinação linear de $v_1, \dots, v_n$, $v_i \in V$, é uma expressão da forma*

$$\alpha_1 v_1 + \cdots + \alpha_n v_n, \tag{2.14}$$

ou, mais sucintamente,

$$\sum \alpha_i v_i. \tag{2.15}$$

- Notação: quando não há ambiguidade, iremos omitir os índices dos somatórios $\sum$.

Definição 5. *Os vetores $v_1, \dots, v_n$ são ditos linearmente independentes (L.I.) se*

$$\sum \alpha_i v_i \iff \alpha_i = 0, \qquad i = 1, \dots, n \tag{2.16}$$

Proposição 2. *$v_1, \dots, v_n$ são L.I. se, e somente se, cada v_i não é o vetor nulo (isto é, $v_i \neq 0$) e nenhum é combinação linear dos demais*

Definição 6. *$\{v_1, \dots, v_n\} \subset V$ é base de V se*

$$\begin{cases} v_1, \dots, v_n \text{ são L.I. e} & (2.17) \\ \text{todo } v \in V \text{ é combinação linear dos } v_i\text{'s} & (2.18) \end{cases}$$

Finalmente, chegamos ao conceito de coordenadas (numéricas) de um vetor:

Definição 7. (Coordenadas) *Dada uma base $B = \{v_1, \dots, v_n\}$ de V, então os números $\alpha_1, \dots, \alpha_n$ tais que que*

$$v = \sum \alpha_i v_i \tag{2.19}$$

são chamados de coordenadas de v em V (na base B). Quando necessário, utiliza-se a notação explícita:

$$\mathcal{X}_B(v) = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}, \tag{2.20}$$

onde o símbolo "chi" $\mathcal{X}$ deve ser lido como "coordenada de".

- Se pensarmos no vetor simbólico $v \in V$ visualmente como uma seta desenhada no plano, suas coordenadas são um vetor numérico de 2 números, sendo necessário definir a base (posição dos eixos x e y).
- Fixada uma base, existe uma relação 1-1 entre vetores simbólicos v e vetores de n números.[1]

Teorema 3. *Dada uma base de n vetores, tem-se que:*

- *As n coordenadas de um vetor são únicas*
- *Qualquer outra base de V tem o mesmo número de elementos n, que usaremos como a definição da dimensão de V.*

Demonstração Ambos decorrem de propriedades de sistemas lineares. Os detalhes já foram vistos pelo aluno no primeiro curso de álgebra linear. ■

- Concluímos que, ao escolher uma base, V pode ser tratado como o espaço numérico K^n (por exemplo, $\mathbb{R}^n$) com as operações numéricas usuais de soma de vetores numéricos e multiplicação de vetor numérico por escalar.

[1] Isto sempre funciona quando a dimensão de V é finita, o que assumiremos sem maiores preocupações.

Aula 3

Matrizes de Transformações

Objetivos

- Representação numérica de transformações lineares: matrizes.
- Conectar álgebra linear simbólica com a numérica, mas a fundo.
- Sem a conexão do numérico com o simbólico, as matrizes perdem significado.
- Como obter os números de uma matriz, a partir do que se deseja realizar na prática?
- Fixar e aprofundar o que já foi visto em outros cursos

Vimos

– revisar conceitos da aula passada

- Vetores como meros símbolos com operações algébricas de soma, e multiplicação por escalar
- Vetores numéricos como coordenadas $\mathcal{X}$ dos vetores simbólicos em uma base.
- V e W irão representar espaços vetoriais

Vamos lembrar a definição simbólica de transformação linear introduzida na Aula 1:

Definição 8. (Simbolicamente Linear) *Uma função $L : V \to W$ é dita linear se*

$$\begin{cases} L(u+v) = L(u) + L(v) & (3.1) \\ L(\alpha v) = \alpha L(v), & (3.2) \end{cases}$$

para todos vetores u, v, e escalares α. Em vez de "função" (vetorial), usamos o termo "transformação" ou "mapa".

Aula de Hoje: Como construir uma matriz

Definição 9. *Dado um mapa $L : V \to W$ e duas bases $A = \{a_1, a_2, \dots, a_n\}$ e $B = \{b_1, b_2, \dots, b_m\}$, a matriz de L relativa às bases A e B é a única matriz $\mathcal{M}_B^A(L)$ tal que:*

$$\boxed{\mathcal{X}_B(L(v)) = \mathcal{M}_B^A(L) \cdot \mathcal{X}_A(v)} \tag{3.3}$$

para qualquer vetor $v \in V$, onde "$\cdot$" é a multiplicação usual de matriz-vetor. Tal matriz é, portanto, uma função entre espaços numéricos de coordenadas $K^n \to K^m$ (por exemplo, de $\mathbb{R}^n \to \mathbb{R}^m$), realizada por multiplicação de matriz usual.

A Equação 3.3 acima é uma das equações mais importantes conectando álgebra linear simbólica com a álgebra linear numérica. O engenheiro deve memorizá-la.

Teorema 4. *As entradas numéricas da matriz $\mathcal{M}_B^A(L)$ são dadas por:*

$$\mathcal{M}_B^A(L) = \begin{bmatrix} | & | & & | \\ M_B^A(L)(\mathcal{X}_A(a_1)) & M_B^A(L)(\mathcal{X}_A(a_2)) & \cdots & M_B^A(L)(\mathcal{X}_A(a_n)) \\ | & | & & | \end{bmatrix} \tag{3.4}$$

$$= \begin{bmatrix} | & | & & | \\ \mathcal{X}_B(L(a_1)) & X_B(L(a_2)) & \cdots & X_B(L(a_n)) \\ | & | & & | \end{bmatrix} \tag{3.5}$$

Ou seja, escreva cada vetor a_i da base A, transformado por L, em termos da base B, e em seguida coloque-os na coluna i da matriz.

Demonstração Como $\mathcal{X}_A(a_i)$ é dado por $(0,\ldots,1,\ldots,0)^\top$, onde 1 ocorre na i-ésima entrada, então a equação 3.3 aplicada a esse vetor resulta na i-ésima coluna da matriz como sendo $\mathcal{X}_B(L(a_i))$. ■

- Implícita na definição de matriz está tanto uma transformação linear (que tem interpretações geométricas), como a escolha de uma base

- É possível, portanto, ter diversas representações numéricas (matrizes) para uma mesma transformação linear.

- A transformação identidade $L = id$, tal que $id(v) = v$, pode ter diversas representações matriciais, além da representação usual com 1's na diagonal.

Definição 10. *(*Matriz de mudança de base*). Dadas duas bases A e B de um mesmo espaço vetorial V, a matriz de mudança de base de A a B é a matriz do mapa identidade relativa às bases A e B, ou seja, $\mathcal{M}_B^A(id)$.*

Observações

- A matriz de mudança de base leva vetores numéricos $\mathcal{X}_A(v)$ a vetores numéricos $\mathcal{X}_B(v)$ via multiplicação usual de matrizes:

$$\mathcal{X}_B(v) = \mathcal{M}_B^A(id)\mathcal{X}_A(v). \tag{3.6}$$

- Apesar da transformação identidade id ser tal que $id(v) = v$, a matriz dessa transformação relativa a bases A e B tal que $A \neq B$ não é a matriz identidade. Em outras palavras, não é o mapa identidade no espaço numérico.

$$\mathcal{M}_B^A(id) : \mathcal{K}^n(\text{coordenadas}) \to \mathcal{K}^n(\text{coordenadas}) \tag{3.7}$$

$$\mathcal{X}_A(v) \mapsto M_B^A(id)\mathcal{X}_A(v) = X_B(v) \tag{3.8}$$

3.0.1 Rotações

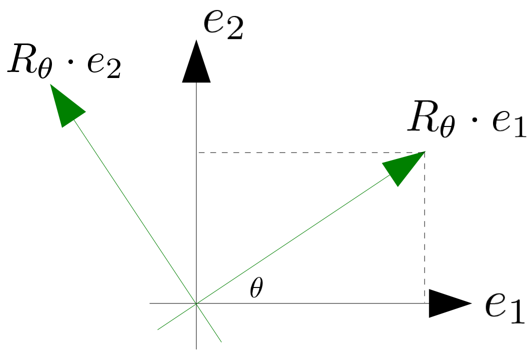
Rotações 2D

Motivação:

- Dada uma rotação (geometria) como definir a matriz de rotação (álgebra)?
- Dada uma mudança de sistemas de coordenadas por uma rotação, como definir a matriz de mudança de base?
- Rotação (geométrica) é linear
  - Se sei como rotacionar base então a rotação está definida (assim funciona para qualquer transformação linear)

Problema 3.1. *Rotacionar um conjunto de pontos 2D no sentido anti-horário em torno da origem.*

Solução.



- Primeiro, o problema geométrico é dotado de uma notação símbólica
- Seja o espaço vetorial o plano 2D ($\mathbb{R}^2$)
- A rotação é uma transformação linear do plano no plano, indicada por Rot
- Para chegarmos a números (afinal, o curso é *numérico*), precisamos especificar coordenadas. Seja o sistema de coordenadas indicado por uma base $E = \{e_1, e_2\}$ e origem O.
- A rotação será especificada por um ângulo θ medido a partir de e_1, no sentido anti-horário.
- Para rotacionar o objeto (conjunto de pontos), precisamos saber como rotacionar um ponto. Em seguida, aplicamos o mesmo procedimento a todos os pontos.

- Pelo diagrama acima, a rotação de um ponto em e_1 é $(\cos\theta, \sin\theta)^\top$, e de um ponto em e_2 é $(-\sin\theta, \cos\theta)^\top$.
- Como vimos, basta colocar esses vetores na coluna da matriz de rotação:

$$R_\theta = \mathcal{M}_E^E(\text{Rot}) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Usar $R(\theta)$

- O sistema de coordenadas não é rotacionado; apenas os pontos do plano são rotacionados, fixando-se o sistema de coordenadas.
- Isso é modelado pela notação simbólica como $\mathcal{M}_E^E(\text{Rot})$, onde a base é fixa, e a transformação é Rot.
- Seguindo-se a notação e o raciocínio de forma cuidadosa, a dúvida usual quanto ao sinal de $\sin\theta$ na matriz desaparece. Controle a ansiedade!
- Outra vantagem é que o mesmo método funciona para qualquer dimensão (3D, 4D, etc)

A seguinte ideia vale para montar a representação numérica de uma matriz de rotação em qualquer dimensão:

Proposição 5. *As* colunas *duma matriz de rotação são a rotação dos vetores base $e_1, \dots, e_n$, escritas nesta própria base. As* linhas *da matriz de rotação são os vetores unitários que são levados aos vetores base* após *a rotação.*

3.0.2 Rotações 3D

Problema 3.2. *Rotacionar uma* base *3D em torno da origem, e escrever a matriz que realiza a mudança de base para um ponto arbitrário.*

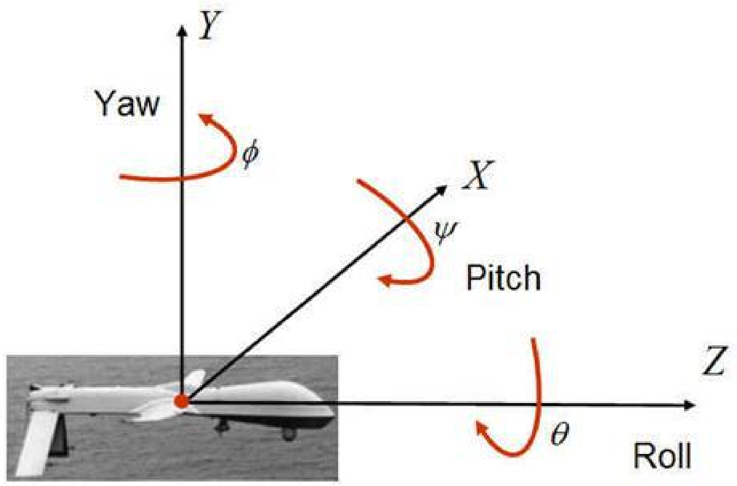
- A primeira dificuldade em definir os números de matrizes de rotação 3D é definir quais ângulos geométricos são positivos (rotações anti-horárias)

Definição 11. *(orientação) Uma base é dita* de mão-direita *quando as rotações por ângulos positivos são tais que, olhando-se de um vetor base em direção à origem, uma rotação anti-horária por $+90^o$ leva um eixo positivo em outro eixo positivo.*

- As regras da mão direita vistas em cursos de física permitem definir na prática a orientação de uma base
- O que nos interessa é a seguinte tabela:

Se o eixo de rotação é	A direção de rotação por ângulo positivo é
x	$y \to z$
y	$z \to x$
z	$x \to y$

Como revisão, a figura abaixo ilustra os graus de liberdade de uma rotação e seus nomes em inglês, no caso de um drone



Em termos de tais ângulos, a matriz de rotação $R = \mathcal{M}_B^A(id)$ (neste caso, de mudança de base, e não de uma rotação geométrica, pois os pontos são fixos) pode ser decomposta como $R(\theta, \phi, \psi) = R_z(\theta)R_y(\phi)R_x(\psi)$, onde

$$R_z(\theta) = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.9}$$

$$R_y(\phi) = \begin{bmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{bmatrix} \tag{3.10}$$

$$R_x(\psi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & \sin\psi \\ 0 & -\sin\psi & \cos\psi \end{bmatrix} \tag{3.11}$$

Note que cada coluna é obtida através da nossa notação:

$$\mathcal{X}_B(v) = \mathcal{M}_B^A(id)\mathcal{X}_A(v), \tag{3.12}$$

aplicada aos vetores base, onde a transformação linear L aqui é id pois estamos fixando os pontos, mas estamos realizando uma mudança de base de A para B.

- A coluna i da matriz é então $\mathcal{X}_B(a_i)$. Em outras palavras, nossa tarefa numérica é escrever cada vetor base a_i *não-rotacionado* em termos dos vetores *rotacionados.*

Exercício 3.1. *Explique detalhadamente como obter a primeira coluna de $R_z(\theta)$.*

A forma final de $R(\theta, \phi, \psi) = R_z(\theta)R_y(\phi)R_x(\psi)$ para uma rotação arbitrária seria:

$$\begin{bmatrix} \cos\theta\cos\phi & \cos\theta\sin\phi\sin\psi + \sin\theta\cos\psi & -\cos\theta\sin\phi\cos\psi + \sin\theta\sin\psi \\ -\sin\theta\cos\phi & -\sin\theta\sin\phi\sin\psi + \cos\theta\cos\psi & \sin\theta\sin\phi\cos\psi + \cos\theta\sin\psi \\ \sin\phi & -\cos\phi\sin\psi & \cos\phi\cos\psi \end{bmatrix} \tag{3.13}$$

Exercício 3.2. *Mostre que esta é a inversa da matriz que realiza a rotação de pontos, mantendo-se a base fixa.*

Aula 4

Modelando Transformações Rígidas Numericamente

Objetivos

- Representação vetorial e numérica de transformações rígidas, nosso primeiro exemplo de transformações não-lineares (translação além de rotação)
- Fornecer exemplos mais práticos para as aulas anteriores
- Qual o poder de modelagem da álgebra matricial? Que tipos de problemas interessantes podemos tratar com essa álgebra simples?

Vimos

– revisar conceitos da aula passada

- Como modelar rotações com matrizes

Aula 5

Coordenadas Homogêneas

Objetivos

- Qual o poder de modelagem da álgebra matricial? Que tipos de problemas interessantes podemos tratar com essa álgebra simples?
- O que podemos fazer apenas com multiplicação matriz-vetor?
- Projeção perspectiva
- Pré-requisito para aulas de sistemas lineares homogêneos mais adiante no curso
- Pré-requisito para exemplo de câmeras, projeção e computação gráfica mais adiante no curso

Vimos

– revisar conceitos da aula passada

- Como modelar rotações e translações numericamente com multiplicação de matriz-vetor seguida de uma soma de vetor

Aula 6

Resolvendo sistemas lineares por SVD

> SVD is the answer.
> What is your problem?
>
> — Joseph Mundy

Objetivos

- Técnica genérica para resolver sistemas lineares
- Técnica mais útil em engenharia para sistemas sobredeterminados por mínimos quadrados usando SVD
- Solução aproximada de equações contraditórias típicas de medições com erro
- Técnica a ser vista não é a mais eficiente, mas é a mais usada antes de algoritmos mais específicos
- Foco será no procedimento geral, sendo a teoria aprofundada adiante

Vimos

– revisar conceitos da aula passada

- Toda matriz $A = V\Lambda U^\top$ (SVD) (ver Aula 1)
- Dado um algoritmo que realize esta decomposição na sua linguagem favorita, como resolver sistemas lineares?
- Não será necessário saber o algoritmo SVD em si, apenas entradas e saídas.

Introdução

- Central em álgebra linear numérica é resolver sistemas da forma
$$Ax = b. \tag{6.1}$$
- Em Scilab ou Matlab, por exemplo, pode-se usar a barra invertida para cegamente tentar obter uma solução:
$$x = A\backslash b, \tag{6.2}$$
onde a notação com a barra invertida "\" sugere que x é b dividido por A, de alguma forma.
- O Scilab nesse caso escolhe o melhor algoritmo, e em geral será o SVD, conforme veremos nesta seção.
- Na prática, a matriz A e o vetor b consistem de medições com erro, repetidas inúmeras vezes.
- Devido a esse erro, não há um x que exatamente satisfaça $Ax = b$.
- Nesse caso, queremos um x que satifaça $Ax = b$ aproximadamente. Ou seja:
$$Ax \approx b, \tag{6.3}$$
onde "$\approx$" aqui significa aproximado.
- A técnica a ser descrita nesta aula é uma maneira prática de usar o SVD.
- Muitas vezes, uma biblioteca em C terá o algoritmo SVD, porém pode não ter outras funções práticas para solução de sistemas lineares.
- Nesta aula, veremos uma técnica que transforma a equação acima numa equação homogênea do tipo $Ax \approx 0$, e em seguida resolve um problema de otimização com SVD.

Homogeneizando Sistemas Lineares

- Dado um sistema $Ax = b$, podemos gerar um sistema equivalente $\tilde{A}\tilde{x} = 0$. Dizemos que "homogeneizamos o sistema".

- O truque tem a ver com coordenadas homogêneas, mas o procedimento a seguir pode ser compreendido mesmo sem lembrar da Aula 5.

- Suponha que a matriz A seja 2×2. O sistema $Ax = b$ ficaria:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \tag{6.4}$$

- Note que apenas b_1 e b_2 não estão acompanhados de variáveis

- Podemos fazer todos os termos conterem variáveis (homogêneos) multiplicando-se todas as equações (em ambos os lados) por uma nova variável w, sem alterar o resultado:

$$\begin{cases} a_{11}x_1w + a_{12}x_2w = b_1w \\ a_{21}x_1w + a_{22}x_2w = b_2w \end{cases} \tag{6.5}$$

  Agora, podemos definir novas variáveis $\tilde{x}_1 = x_1w$, $\tilde{x}_2 = x_2w$, de forma que nosso novo vetor $\tilde{x}$ fique:

$$\begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \end{bmatrix} \doteq \begin{bmatrix} wx_1 \\ wx_2 \\ w \end{bmatrix} = w \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}. \tag{6.6}$$

  Assim:

$$\begin{bmatrix} a_{11} & a_{12} & -b_1 \\ a_{21} & a_{22} & -b_2 \end{bmatrix} \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \end{bmatrix} = 0 \tag{6.7}$$

- Tal equação é homogênea:

$$\tilde{A}\tilde{x} = 0 \tag{6.8}$$

  A solução será o Kernel de A, $\ker A$ (núcleo de A).

- Logo, usando coordenadas homogêneas, podemos reduzir qualquer sistema linear não-homogêneo a um sistema linear homgoêneo

- Torna-se central, então, saber resolver:

$$Ax = 0, \qquad \text{para } |x| \neq 0 \tag{6.9}$$

- Por exemplo, exigindo $|x| = 1$
- Uma vez encontrado algum elemento $\tilde{x}$ do espaço de soluçao $\ker \tilde{A}$, para um sistema $\tilde{A}\tilde{x} = 0$, podemos obter a solução para os sistema original $Ax = b$
- Basta normalizar $\tilde{x}$ para obter $w = 1$ (sistema original), o que ocorre fazendo-se:

$$x = \begin{bmatrix} \tilde{x}_1/\tilde{x}_3 \\ \tilde{x}_2/\tilde{x}_3 \end{bmatrix} \tag{6.10}$$

Solução aproximada por SVD: método prático

Resolver o sistema aproximado

$$Ax \approx 0, \qquad \text{e } |x| = 1, \tag{6.11}$$

é o mesmo que exigir que Ax é pequeno. Ou seja, queremos resolver o seguinte problema de otimização

$$\arg\min_{|x|=1} |Ax|. \tag{6.12}$$

- Se imaginarmos x como um círculo ou esfera unitária, $|Ax|$ será mínimo na direção do vetor singular correspondente ao menor valor singular, ou seja, basta calcular a SVD. (isto foi visto intuitivamente nas aulas anteriores)
- Algebricamente, temos que, pelo SVD de A:

$$|Ax| = |V\Lambda U^\top x| = |\Lambda U^\top x|, \tag{6.13}$$

onde a última passagem se deve ao fato de V ser ortogonal, ou seja, não altera a norma.

- Logo, queremos minimizar

$$|\Lambda U^\top x| \quad \text{tal que } |x| = 1 \tag{6.14}$$

- Como $|x| = |U^\top x|$, podemos definir $y \doteq U^\top x$.
- Então, queremos minimizar

$$|\Lambda y| \quad \text{tal que } |y| = 1. \tag{6.15}$$

- Como Λ é diagonal, e os algoritmos retornam os valores singulares σ_i na ordem decrescente, então a solução é

$$y = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \tag{6.16}$$

- Dessa forma, tínhamos

$$y = U^\top x \implies x = Uy \tag{6.17}$$

- Logo x é a última coluna de U na decomposição SVD de A.

O método padrão SVD para resolver $Ax \approx b$ para n variáveis $x_1, \ldots, x_n$ pode ser resumido da seguinte forma

1. Homogeneiza-se o sistema para $\tilde{A}\tilde{x} \approx 0$, que terá $n+1$ variaveis
2. Calcula-se o SVD de $\tilde{A}$
3. A última coluna de U é o $\tilde{x}$ que resolve $\tilde{A}\tilde{x} \approx 0$
4. Obtém-se x a partir de $\tilde{x}$ ignorando-se a última coordenada do vetor normalizado $\tilde{x}/\tilde{x}_{n+1}$

Aula 7

Aprofundando em SVD

Objetivos

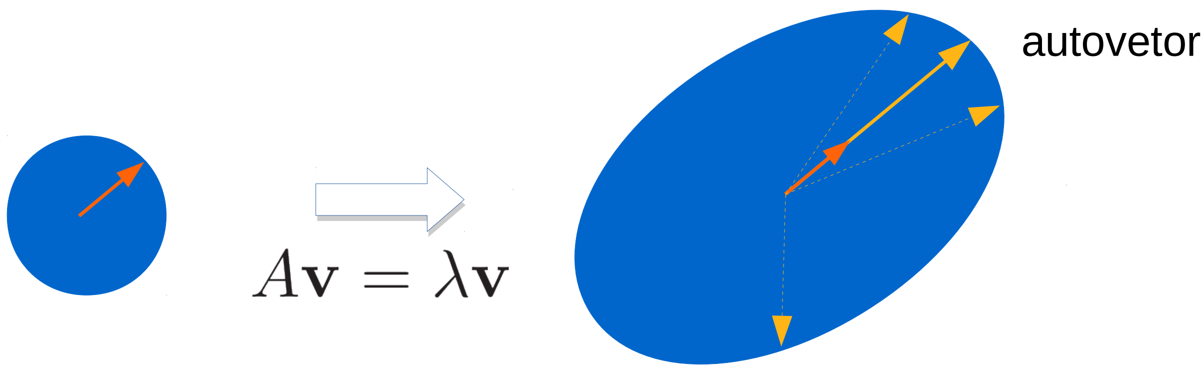
- Definição formal de SVD após ter visto como usar
- Demonstração do SVD – um dos teoremas mais importantes da álgebra linear moderna, por causa de um algoritmo de álgebra lihnear numérica.
- Conexão com autovalores e autovetores
- Gancho para análise de algoritmos para SVD e para autovalores e autovetores do Golub (este curso é estilo *top-bottom*)
- Interpretações gráficas de autovalores, autovetores, valores singulares e vetores singulares

Vimos

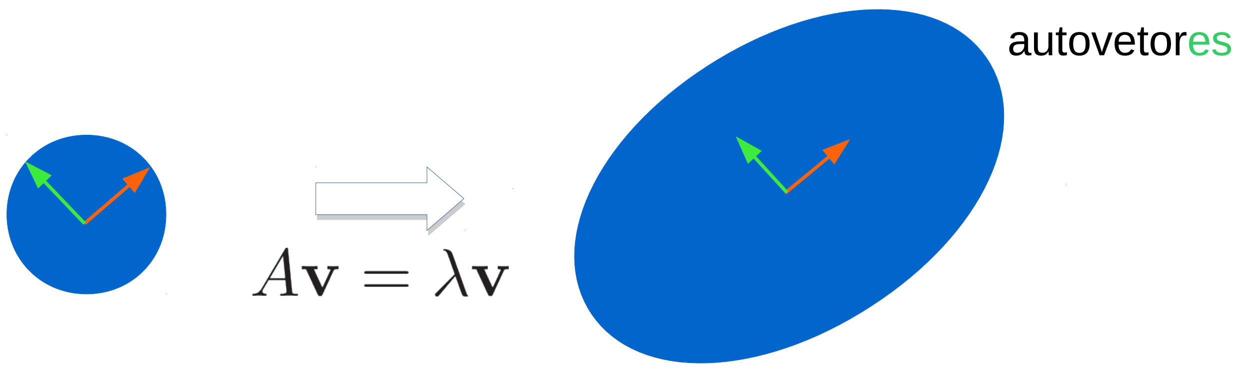
– revisar conceitos da aula passada

- Já vimos SVD operacionalmente: $A = U\Lambda V^\top$
- Já vimos como SVD pode ser usado para resolver sistemas $Ax \approx b$ de forma prática, mesmo não sendo a mais eficiente em todos os casos.
- Veremos agora a definição formal de SVD

- Um autovetor de trasnformação linear é um vetor cuja direção é preservada pela transformação linear:



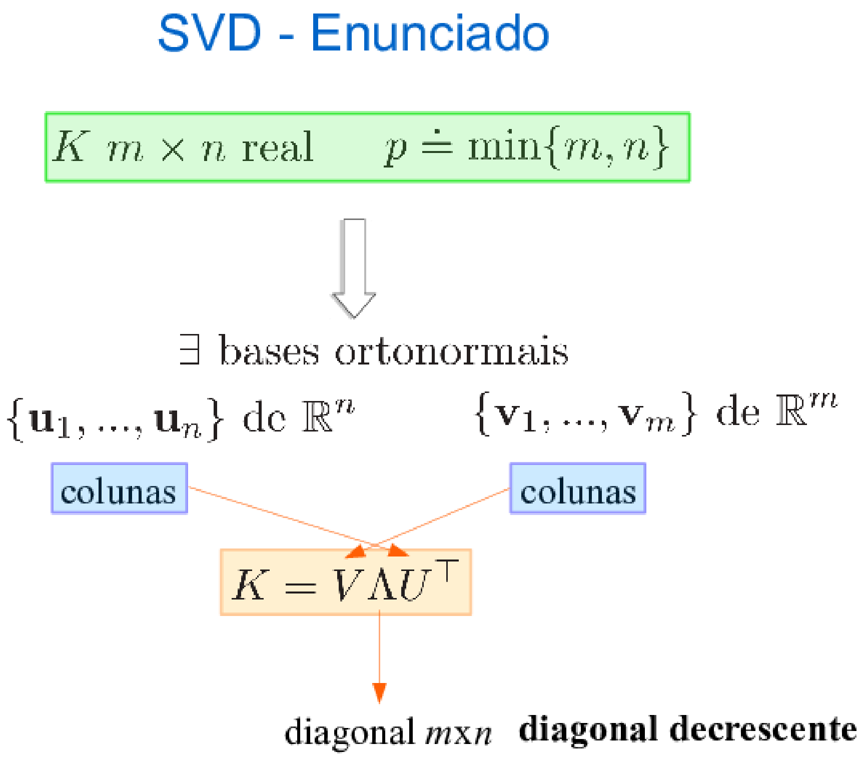
- A figura acima mostra que a matriz A estica o disco a uma elipse, uma manifestação numérica de uma transformação linear
- Se imaginarmos que o disco é feito de borracha, e a matriz A é tal que o esticamento é feito puxando-se os pontos do disco original ao longo de eixos da elipse, então a direção dos pontos ao longo dos eixos não se altera, pois os pontos apenas movem ao longo dos eixos durante o esticamento
- Para tais tipos de matrizes, podemos tomar todos os eixos da elipse como uma base:



Nessa base, *A* fica diagonal → diagonalização

- Como fica esse raciocínio para *qualquer* matriz K de tamanho $m \times n$?

- A resposta é SVD. Veja um esquema do enunciado que veremos formalmente a seguir.



- Note que, diferenciar de A acima, apropriada para autovetores e autovalores, iremos utilizar a letra K para uma matriz geral.

Teorema 6. (Decomposição em Valores Singulares) *Seja K uma matriz real $m \times n$ e $p \doteq \min\{m, n\}$ Então existem bases ortonormais $\{u_1, \ldots, u_n\}$ de $\mathbb{R}^n$ e $\{v_1, \ldots, v_m\}$ de $\mathbb{R}^m$ tais que*

$$K = V\Lambda U^\top, \tag{7.1}$$

com Λ matriz diagonal $m \times n$ com elementos diagonais

$$\Lambda = \begin{bmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \sigma_n \\ 0 & & \ldots & 0 \\ \vdots & & & \vdots \\ 0 & & \ldots & 0 \end{bmatrix} \tag{7.2}$$

, se $m > n = p$, ou

$$\Lambda = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & \ddots & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & & \vdots \\ 0 & \dots & 0 & \sigma_{\boldsymbol{m}} & 0 & \dots & 0 \end{bmatrix} \tag{7.3}$$

, se $p = m < n$, e as matrizes

$$V = \begin{bmatrix} | & & | \\ v_1 & \dots & v_m \\ | & & | \end{bmatrix}_{m\times m} \qquad e \quad U = \begin{bmatrix} | & & | \\ u_1 & \dots & u_n \\ | & & | \end{bmatrix}_{n\times n}$$

são ambas ortogonais. Os números $\sigma_1, \dots, \sigma_p$ são chamados valores singulares de K.

Demonstração.

- Note que K define uma transformação linear $\mathbb{R}^n \to \mathbb{R}^m$.
- Note-se, também, que K é arbitrária. A idéia do teorema é montar uma matriz simétrica, adequada para o cálculo de autovalores e autovetores. Dessa forma, o SVD é reduzido a autovalroes e autovetores.
- Há duas formas óbvias de montar uma matriz simétrica a partir de K:
- A primeira forma é $K = K^\top K$, uma matriz simétrica de tamanho $n \times n$.
- A segunda é $K = KK^\top$, uma matriz simétrica de tamanho $m \times m$.
- Os autovetores dessas matrizes dão os vetores singulares
- Os autovalores dessas matrizes são iguais e dão os valores singulares ao quadrado
- A seguir, vamos detalhar esses passos de uma forma menos redundante

1. Monta-se uma base ortonormal para $\mathbb{R}^n$ (domínio)

- Seja $A = K^\top K$, a qual é real e simétrica, satisfazendo as hipóteses do teorema espectral
- Sejam $u_1, \dots, u_n$ os vetores da base ortonormal de $\mathbb{R}^n$ que diagonaliza A

- Sejam $\lambda_1 \geq \cdots \geq \lambda_n$ tal que

$$Au_i = \lambda u_i, \quad i = 1, \ldots, n. \tag{7.4}$$

2. Monta-se uma base ortonormal para $\mathbb{R}^m$ (contra-domínio) Seja n_0 o índice do menor λ_i e

$$w_i \doteq Ku_i, \quad i = 1, \ldots, n_0 \leq p. \tag{7.5}$$

Então os w_i são ortogonais:

$$w_i^\top w_j = u_i^\top K^\top K u_j = u_i^\top A u_j = \lambda_j u_i^\top u_j = \lambda_j \delta_{ij}, \tag{7.6}$$

onde δ_{ij} é o delta de Kronecker (1, se $i = j$, 0 se $i \neq j$) para $i, j = 1, \ldots, n_0$

Definem-se

$$v_i \doteq \frac{w_i}{|w_i|} = \frac{Ku_i}{|Ku_i|}, \quad i = 1, \ldots, n_0, \tag{7.7}$$

$$\sigma_i \doteq |Ku_i|. \tag{7.8}$$

Logo, $v_1, \ldots, v_{n_0}$ são ortonormais e

$$Ku_i = \sigma_i v_i \tag{7.9}$$

e

$$K^\top v_i = \frac{K^\top K}{\sigma_i} u_i = \frac{\lambda_i}{\sigma_i} u_i. \tag{7.10}$$

Ademais, $\lambda_i = \sigma_i^2$, $i = 1, \ldots, n_0$, pois

$$\lambda_1 = \lambda_i u_i^\top u_i = (\lambda_i u_i)^\top u_i = (Au_i)^\top u_i \tag{7.11}$$

$$= (K^\top K u_i)^\top u_i = u_i^\top K^\top K u_i = (Ku_i)^\top (Ku_i) \tag{7.12}$$

$$= \sigma_i^2 V_i^\top V_i = \sigma_i^2, \quad i = 1, \ldots, n_0 \tag{7.13}$$

- Definimos $\sigma_i \doteq 0$, para $n_0 < i \leq p$, se existir tal i

- Sejam v_i, $n_0 < i \leq m$, vetores ortonormais completando $v_1, \ldots, v_{n_0}$ a uma base ortonormal de $\mathbb{R}^m$

- Montando-se as matrizes
$$V = \begin{bmatrix} | & & | \\ v_1 & \ldots & v_m \\ | & & | \end{bmatrix}_{m\times m} \quad \text{e} \quad U = \begin{bmatrix} | & & | \\ u_1 & \ldots & u_n \\ | & & | \end{bmatrix}_{n\times n},$$
as propriedades enunciadas são satisfeitas.

■

Discussão 1

- Dada a decomposiçào SVD $K = V\Lambda U^\top$, multiplicando-se ambos os lados por U, tem-se:
$$KU = V\Lambda, \tag{7.14}$$
ou seja
$$Ku_i = \sigma_i v_i, \tag{7.15}$$
uma equação análoga à definição de autovalores e autovetores.
- Na equação, parece ser arbitrário que ao se multiplicar um vetor, obtém-se o múltiplo de outro vetor
- É importante que u_i's e v_i's definem bases ortogonais para $\mathbb{R}^n$ e $\mathbb{R}^m$, respectivamente

Discussão 2

- Se $K = V\Lambda U^\top$, então $K^\top K = U\Lambda V^\top V\Lambda U^\top = U\Lambda^2 U^\top$, uma diagonalização comum com autovetores e autovalores
- E $KK^\top = V\Lambda^2 V^\top$, uma diagonalização comum com autovetores e
- O SVD fornece bases ortonormais aos seguintes espaços fundamentais de K:
$$\begin{cases} v_1, \ldots, v_r \text{ base ortonormal de } \operatorname{Im}(K) & (7.16) \\ v_{r+1}, \ldots, v_m \text{ base ortonormal de } \ker(K^\top) & (7.17) \\ u_1, \ldots, u_r \text{ base ortonormal de } \operatorname{Im}(K^\top) & (7.18) \\ u_{r+1}, \ldots, u_n \text{ base ortonormal de } \ker(K) & (7.19) \end{cases}$$

[todo: propriedade da inversa por SVD]

Aula 8

Decomposição QR

Objetivos

- Principais algoritmos de decomposição QR

Vimos

– revisar conceitos da aula passada

- Vimos a decomposição SVD.
- A decomposição QR é um passo importante em muitos outros algoritmos, além de ser melhor em casos específicos

8.1 Revisão de matrizes ortogonais

- Uma matriz quadrada U é ortogonal se $U^\top U = UU^\top = I$
- Propriedade: $\det U = \pm 1$
- Matrizes de rotação são matrizes ortogonais com determinante $+1$
- Matrizes ortogonais preservam produto interno e norma: $(Ux)^\top Ux = x^\top U^\top Ux = x^\top x$

8.2 Decomposiçào QR por matrizes Givens

Definição 12. *A decomposição QR de uma matriz A é da forma:*

$$A = QR$$

Ortogonal

Triangular superior
“Right-triangular”

- Existem também decomposições similares à QR, todas chamadas “QR” em um sentido amplo: QL, LQ, RQ, onde “L” é triangular inferior (de “left / lower-triangular”).
- Antes de entrar em detalhes tétnicos, vejamos um algoritmo simples para matrizes 3×3 usando matrizes especiais chamadas “Givens”

Definição 13. *Uma matriz Givens de rotação é uma rotação em torno de um dos eixos x, y ou z:*

$$Q_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{bmatrix}, \quad Q_y = \begin{bmatrix} c & 0 & s \\ 0 & 1 & 0 \\ -s & 0 & c \end{bmatrix}, \quad Q_z = \begin{bmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{8.1}$$

, onde $c = \cos\theta$ e $s = sen\theta$.

- Ao fazer a operação $A \cdot Q_z$ para uma matriz A, sua última coluna não muda, e as duas primeiras é combinação linear das duas primeiras de A.
- (multiplicar AQ_z explicitamente para constatar isso)
- Podemos escolher θ de forma a zerar qualquer entrada desejada das primeiras colunas
- Por exemplo, para setar o elemento (2,1) do resultado para zero,

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} ca_{11} + sa_{12} & -sa_{11} + ca_{12} & a_{13} \\ ca_{21} + sa_{22} & -sa_{21} + ca_{22} & a_{23} \\ ca_{31} + sa_{32} & -sa_{31} + ca_{32} & a_{33} \end{bmatrix}. \tag{8.2}$$

Impondo

$$\begin{cases} ca_{21} + sa_{22} = 0 & (8.3) \\ c^2 + s^2 = 1, & (8.4) \end{cases}$$

encontramos c e s.

8.2.1 Estratégia de Algoritmo RQ por matrizes Givens

- Iremos descrever uma estratégia para o algoritmo RQ, sendo o QR análogo.
- Zerar a parte inferior da matriz, uma entrada por vez, multiplicando-se por matrizes Givens, cuidando-se para não modificar as entradas que já foram anuladas

---

Algorithm 1 Decomposição RQ por matrizes Givens

Entrada: matriz 3×3 A
Saída: R e Q tal que $A = RQ$

início

1. Multiplica-se por Q_x tal que $A_{32} \to 0$
2. Multiplica-se por Q_y tal que $A_{31} \to 0$ (isso não muda A_{32} pois não afeta a coluna 2
3. Multiplica-se por Q_z tal que $A_{21} \to 0$ (A_{31} e A_{32} permanecem zeradas, já que as primeiras colunas são substituídas por combinações lineares delas mesmas)
4. Observação: outras sequências de rotações Givens podem também resultar na mesma coisa
5. $AQ_xQ_yQ_z = R \implies A = RQ_z^\top Q_y^\top Q_x^\top = RQ$

fim

Aula 9

Matrizes Householder para Decomposição QR

Objetivos

- Ilustar o importante assunto de Matrizes Householder em álgebra linear numérica

Vimos

– revisar conceitos da aula passada

- Vimos a decomposição QR por matrizes Givens.

- Para matrizes grandes $n \times n$, com $n \gg 3$, é melhor usar matrizes de Householder, que são da forma:

$$H_v = I - 2\frac{vv^\top}{v^\top v} \tag{9.1}$$

- H_v aplicada a um vetor a é:

$$H_v a = (I - 2\frac{vv^\top}{v^\top v})a = a - 2\frac{v(v^\top a)}{v^\top v} \tag{9.2}$$

- Note que H_v é ortogonal e simétrica (mostrar em aula)

Proposição 7. *Seja*

$$v = x \pm |x| e_1, \tag{9.3}$$

onde e_1 *é um vetor base. Então*

$$H_v x = \mp |x| e_1. \tag{9.4}$$

Ou seja, H_v *é uma matriz ortogonal que leva* x *a um múltiplo de* e_1*, portanto zerando todas as outras entradas / coordenadas.*

Demonstração.

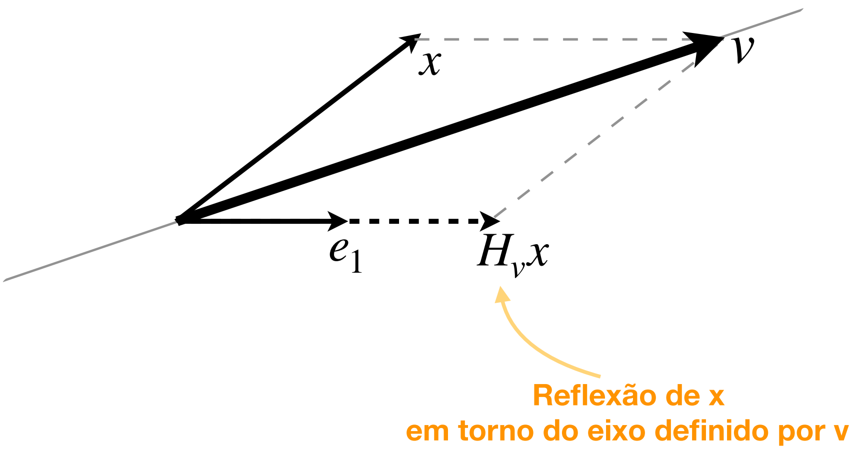
$$H_v x = I - 2\frac{(x \pm |x|e_1)(x \pm |x|e_1)^\top x}{(x \pm |x|e_1)^\top (x \pm |x|e_1)} \tag{9.5}$$

- Exercício em aula: terminar esta prova

■

Toma-se

$$v = x + sign(x_1)|x|e_1 \tag{9.6}$$



Para uma dada matriz A, tomando-se x sua primeira coluna e v como acima, então $H_v A$ vai zerar quase toda primeira coluna, que fica

$$\begin{bmatrix} |x| \\ \vdots \\ 0 \end{bmatrix}. \tag{9.7}$$

Continuando a multiplicação à esquerda por rotações Householder adequadas, podemos triangular a matriz:

$$\tilde{Q}A = R \tag{9.8}$$
$$A = \tilde{Q}^\top R = QR \tag{9.9}$$

que é a decomposição QR de A.

Tais matrizes Householder adequadas são da seguinte forma:

$$Q_1 = H_v, \quad \text{com } v \text{ definido em relação a } x \text{ sendo a primeira coluna de } A$$
$$Q_1 A = \left(\begin{array}{c|ccc} \alpha_1 & \bullet & \dots & \bullet \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array}\right) \tag{9.10}$$

- Delete a primeira linha e a primeira coluna de $Q_1 A$ para obter A'
- Aplica-se a matriz Householder para A', obtendo-se Q'_2
- Faça-a operar em A expandindo-se com 1's na diagonal:

$$\begin{bmatrix} I_{k-1} & 0 \\ 0 & Q'_k \end{bmatrix} \tag{9.11}$$

- Assim:

$$R = Q_t \dots Q_2 Q_1 A, \tag{9.12}$$
$$A = QR, Q \qquad = Q_1^\top Q_2^\top \dots Q_t^\top \tag{9.13}$$

Exercício 9.1. *(em sala de aula) Realize a decomposição QR da seguintes matriz via matriz Givens e via Householder:*

$$\begin{bmatrix} 12 & 91 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{bmatrix} \tag{9.14}$$

Aula 10

Decomposição de Cholesky

Objetivos

- Compreender aspectos básicos da decomposição de cholesky

Vimos

– revisar conceitos da aula passada

- Vimos as decomposições QR e SVD.
- Na demonstração do SVD, tinhamos matrizes simétricas A montadas a partir de K: $A = K^\top K$ ou $A = KK^\top$
- Em algumas aplicações, temos o contrário: dada uma matriz simétrica A, como encontrar K tal que $A = K^\top K$ ou $A = KK^\top$
- A decomposição de Cholesky é um tipo de raíz quadrada de matrizes
- É uma raíz quadrada útil quando a matriz é simétrica, positiva definida (revisaremos a definição)
  - Muitos casos práticos são desse tipo
  - Nesses casos, Cholesky é bem mais rápido e estável que alternativas
- Esta decomposição é parte do ferramental fundamental de álgebra linear numérica

- Dada uma matriz A, a decomposição de cholesky retorna K tal que $A = KK^\top$, se possível
- Se A fosse um número, seria exigido que A fosse positivo (e real), caso quisessemos uma raíz quadrada K real
- Similarmente, para uma matriz A quadrada $n \times n$, exigimos que seja:

Condição na matriz	**Condição correspondente nos autovalores**
Simétrica	autovalores reais
Positiva-definida	autovalores positivos

Definição 14. *A é simétrica positiva-definida se todos seus autovalores são positivos*

Teorema 8. *Os seguintes fatos são verdadeiros:*

1. *Uma matriz simétrica A é positiva-definida $\iff x^\top Ax > 0, \forall x \neq 0$*
2. *Uma matriz A é simétrica positiva definida $\implies$ existe uma única matriz K triangular superior e real com diagonal positiva tal que $A = KK^\top$*

Demonstração.

1. (a) Por decomposição em autovalores e autovetores, temos $A = UDU^\top$

   (b) Ida:
   $$x^\top UDU^\top x = yDy^\top = \sum d_{ii} y_i^2 > 0, \forall x \tag{10.1}$$

   (c) Volta:
   $$x^\top Ax > 0 \implies x^\top UDU^\top x > 0 \implies y^\top Dy > 0, \forall y \text{ onde } y = x^\top U. \tag{10.2}$$
   Colocando $y = e_i$, temos $d_{ii} > 0$

2. - $A = UDU^\top$, com D real positiva, então tomemos a "raiz quadrada de D",
     $$D = EE^\top, \quad \text{com } E \text{ diagonal.} \tag{10.3}$$
   - Logo, $A = VV^\top$, onde $V = UE$.

- A matriz V não é necessariamente triangular superior.
- Aplica-se, então, a decomposição RQ da forma

$$V = KQ, \tag{10.4}$$

o que dá

$$A = VV^\top = KQQ^\top K^\top = KK^\top \tag{10.5}$$

- Se as diagonais não forem positivas, basta multiplicar K à direita por uma matriz com ± 1's nas diagonais – isso não afetará o produto $K^\top K$
- K é única pois:
- Sejam K_1, K_2 com $K_1 K_1^\top = K_2 K_2^\top$. Então

$$K_2^{-1} K = K_2^\top K_1^{-\top} = D, \tag{10.6}$$

duas matrizes tiangualres superiores e inferiores, respectivamente.
- Ademais, $K_2^{-1} K_1$ é igual a $(K_2^\top K_1^{-\top})^{-\top}$
- Logo, $D = D^{-\top}$ sendo, portanto, com entradas ± 1. Como K_1 e K_2 têm diagonal positiva, este sinal é positivo.

■

Vejamos um primeiro algoritmo para decomposição de Cholesky de matrizes 3×3:

Algorithm 2 Algoritmo de Cholesky-Banachiewicz

Entrada: matriz 3×3 A

Saída: matriz 3×3 K tal que $A = KK^\top$

início

1. Suponha K triangular inferior L e multiplique as matrizes:

$$A = LL^\top = \begin{bmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & 0 \\ L_{31} & L_{32} & L_{33} \end{bmatrix} \begin{bmatrix} L_{11} & L_{21} & L_{31} \\ 0 & L_{22} & L_{32} \\ 0 & 0 & L_{33} \end{bmatrix} = \tag{10.7}$$

$$= \begin{bmatrix} L_{11}^2 & \bullet & \bullet & \\ L_{21}L_{11} & L_{21}^2 + L_{22}^2 & \bullet & \\ L_{31}L_{11} & L_{31}L_{21} + L_{32}L_{22} & L_{31}^2 + L_{31}^2 + L_{33}^2 & L_{33} \end{bmatrix}. \tag{10.8}$$

   Onde os símbolos "$\bullet$" denotam entradas simétricas.

2. Resolve-se os L_{ij}'s em termos de A_{ij} linha por linha, iniciando pelo canto superior esquerdo:

$$L_{jj} = \sqrt{A_{jj} - \sum_{k=1}^{j-1} L_{jk}^2} > 0, \quad \text{se } A \text{ é positiva definida} \tag{10.9}$$

$$L_{ij} = \frac{1}{L_{jj}} \left(A_{ij} - \sum_{k=1} j - 1 L_{ik} L_{jk} \right), \qquad \text{para } i > j \tag{10.10}$$

fim

Aula 11

Revisão de Scilab/Matlab/Python

Objetivos

- Introduzir resumo de aspectos matriciais de linguagem de programação *script*, visando realizaçao das tarefas
- Geralmente dada no início do curso. Fica a critério do instrutor quando dar esta aula.

Aula 12

Exponenciais de Matrizes

Objetivos

- Métodos numéricos para solução de equações não-lineares exigem a solução de uma família de sistemas lineares
- Faz-se necessário obter essa família a partir de parametrizações simples
- Séries de potência permitem gerar famílias de sistemas lineares expandindo a partir de matrizes simples.
- A expansão em série pode ser uma base para algoritmos eficientes, mesmo não sendo em si um procedimento final
- Base para aplicações em sistemas dinâmicos e solução de equações não-lineares usando teração com sistemas lineares (métodos de Newton, Gauss-Newton e afins).

Vimos

– revisar conceitos da aula passada

12.1 Matrizes Antissimétricas

- Será útil revisarmos matrizes antissimétricas e aprofundar neste assunto

- Na Seção 8.1, revisamos matrizes ortogonais que serão úteis como exemplos nesta seção.
- Resta-nos revisar matrizes antissimétricas
- Matrizes antissimétricas e ortogonais serão conectadas com exponenciais de matrizes
- Métodos numéricos relacionados serão obtidos e analisados, permitindo o uso de matrizes simples para expressar e parametrizar matrizes mais complicadas.
- Dado um vetor $\mathbf{v} = (v_1, v_2, v_3)^\top$, é possível construir uma matriz antissimétrica com as componentes de $\mathbf{v}$ na forma

$$[\mathbf{v}]_\times = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix}.$$

- A notação $[\mathbf{v}]_\times$, também denotada $\mathbf{v}_\times$, indica uma relação do produto desta matriz por um vetor qualquer $\mathbf{u}$ com o produto vetorial entre $\mathbf{u}$ e $\mathbf{v}$. De fato, pode-se expressar o produto vetorial por qualquer vetor $\mathbf{v}$ através de multiplicação de matriz:

$$\mathbf{v} \times \mathbf{u} = [\mathbf{v}]_\times \mathbf{u} = (\mathbf{v}^\top [\mathbf{u}]_\times)^\top$$

- Ademais, para toda matriz antissimétrica A existe um vetor $\mathbf{v}$ tal que $A = \mathbf{v}_\times$.
- Tal representação matricial de um produto vetorial e a interpretação da multiplicação por matriz antissimétrica como tal tem grandes implicações práticas para o uso de técnicas matriciais na solução de equações envolvendo 3 dimensões (mecânica/cinemática).
- O produto vetorial entre um vetor e ele mesmo é sempre o vetor nulo
- Portanto o vetor $\mathbf{v}$ é o vetor nulo à direita e à esquerda de $[\mathbf{v}]_\times$. Ou seja, $\mathbf{v}_\times \mathbf{u} = 0$ e $\mathbf{u}^\top \mathbf{v}_\times$.
- Em outras palavras, as linhas e colunas de $\mathbf{v}_\times$ são ortogonais a $\mathbf{v}$.

- O posto de uma matriz antissimétrica $\mathbf{v}_\times$ é 2 se $v \neq 0$
- Desta forma, uma matriz antissimétrica 3×3 será sempre definida por seu vetor nulo. Isto pode ser mostrado para qualquer $n \times n$ ímpar usando determinantes.
- O conjunto das matrizes antissimétricas $n \times n$ juntoamente com operação de multiplicação matriz-matriz é denotado $so(n)$.
- Uma matriz antissimétrica qualquer M satisfaz a relação $\mathbf{v}^\top M\,\mathbf{v} = 0$:

$$
\begin{aligned}
\mathbf{v}^\top M\,\mathbf{v} &= \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} \begin{bmatrix} 0 & m_{12} & m_{13} \\ -m_{12} & 0 & m_{23} \\ -m_{13} & -m_{23} & 0 \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \\
&= \begin{pmatrix} -v_2m_{12} - v_3m_{13} & v_1m_{12} - v_3m_{23} & v_1m_{13} + v_2m_{23} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \\
&= -v_1v_2m_{12} - v_1v_3m_{13} + v_1v_2m_{12} - v_2v_3m_{23} + v_1v_3m_{13} + v_2v_3m_{23} \\
&= 0.
\end{aligned}
$$

Ademais:

- Se $\mathbf{u}$ é tal que $\|\mathbf{u}\| = 1$ e $U := \mathbf{u}_\times$, então

$$U^2 = uu^\top - I \text{ e } U^3 = -U. \tag{12.1}$$

  Se $\|\mathbf{u}\| \neq 1$, a última relação decorre de $U^3 = -\|u\|^2 U = 0$.

- Toda matriz é a soma de uma matriz simétrica e uma antissimétrica:

$$A = \frac{1}{2}\left(A - A^\top\right) + \frac{1}{2}\left(A + A^\top\right). \tag{12.2}$$

- (Transformação de Cayley) Dada uma matriz antissimétrica A então:

  1. Vale a seguinte identidade para A antissimétrica:

$$(I + A)^{-1}(I - A) = (I - A)(I + A)^{-1} \tag{12.3}$$

  2. Tal matriz é ortogonal e, de fato, uma rotação

3. Veremos que matrizes antissimétrics (produtos vetoriais em 3D) podem ser interpretadas como rotações por um ângulo suficientemente pequeno.
4. Exemplo: multiplicando-se uma matriz específica por um vetor 2D, depois 3D, ve-se o efeito.

[todo: figura com rotação local]

Definição 15. *O comutador* $[A, B]$ *de duas matrizes* A *e* B *é a diferença entre multiplicar na ordem* AB *e na ordem* BA*:*

$$[A, B] \doteq AB - BA.$$

- Nota-se que o comutador de matrizes antisimétricas é também uma matriz antissimétrica: $[A, B] = -[B, A]$.
- Todos os auto-valores de A antissimétrica são ou zero ou puramente imaginários, da forma $i\omega$ para algum $\omega \in \mathbb{R}$.
- Isto é intuitivo, já que A permuta coordenadas com sinal trocado, logo nenhum vetor será múltiplo dele mesmo; também é fácil de ver a partir da interpretação por produto vetorial.

Teorema 9. *(Decomposição de matrizes antissimétricas) [?] Toda matriz* A *antissimétrica pode ser decomposta da forma:*

$$A = V \Lambda V^\top, \tag{12.4}$$

onde Λ *é uma matriz diagonal por blocos* $\Lambda = diag\{A_1, \dots, A_m, 0, \dots, 0\}$*, onde cada* A_i *é uma matriz antissimétrica real* 2×2 *da forma:*

$$A_i = \begin{bmatrix} 0 & a_i \\ -a_i & 0 \end{bmatrix}, \quad i = 1, \dots, m. \tag{12.5}$$

- Se A é qualquer, então $A^\top \mathbf{u}_\times A$ é trivialmente antissimétrica
- Logo, existe $\mathbf{v}$ tal que $A^\top \mathbf{u}_\times A = \mathbf{v}_\times$.
- Vamos tentar encontrar uma fórmula para $\mathbf{v}_\times$ a partir de A e u a seguir.

Proposição 10. *Se A é uma matriz 3×3 de determinante 1, então*

$$A^\top \mathbf{u}_\times A = \left[A^{-1}\mathbf{u}\right]_\times \tag{12.6}$$

- A proposição 10 permite "trocar de lado" uma matriz A com uma antissimétrica da seguinte forma:

$$\mathbf{u}_\times A = A^{-\top} A^\top \mathbf{u}_\times A = A^{-\top}\left[A^{-1}\mathbf{u}\right]_\times \tag{12.7}$$

Exercício 12.1. *Como fica o resultado da proposição 10 se A é não-inversível ou com determinante não-unitário?*

Exponenciais de Matrizes

- A função exponencial pode ser aplicada a matrizes tratando-se as matrizes como números na série de potências e usando multiplicação de matrizes.

$$R = \exp(A) = e^A := \sum_{k=0}^{\infty} \frac{A^k}{k!} = I + \sum_{k=1}^{\infty} \frac{A^k}{k!} = I + A + \frac{A^2}{2} + \frac{A^3}{6} + \frac{A^4}{24} + \cdots \tag{12.8}$$

- Isso funciona mesmo com o produto de matrizes sendo não-comutativo, mas por esta não-comutatividade as propriedades desta função exponencial tem diferenças com o caso de números.

- A função exponencial é um exemplo de uma função de matrizes, e é uma das funções de matrizes mais utilizadas em computação [?].

- Exercício: se A é antissimétrica, $exp(A)$ é uma matriz ortogonal.

- Exercicio: Ademais, $det(exp(A)) = +1$, então e^A é uma rotação.

Teorema 11. *(Fórmula de Rodrigues) Para uma matriz anti-simétrica $A = \mathbf{v}_\times$, a matriz de rotação R pode ser calculada usando uma simplificação da série da função exponencial*

$$R = e^A := I + \frac{A}{\|v\|}\sin(\|v\|) + \frac{A^2}{\|v\|^2}(1 - \cos(\|v\|)), \tag{12.9}$$

a qual permite o cálculo usando apenas duas potências da matriz, e funções trigonométricas de números reais.

Demonstração. Usando as identidades (12.1), re-escreva a série de potência da exponencial de matrizes em termos da séries de Taylor de seno e cosseno, separando-se as potências ímpares das pares. ∎

- A partir deste momento, após dar alguns exemplos para montar a intuição do aluno, o instrutor pode dar uma ou duas aulas acerca da análise de estabilidade do cálculo da exponencial de matrizes dado no Golub, procurando simplificar a analise. Isto, acompanhado de revisão do cálculo do mesmo algoritmo e de raízes para números reais.

- Este capítulo aborda um aspecto mais detalhado da álgebra linear numérica, o que permite ao aluno comparar técnicas numéricas versus matriciais para problemas de avaliação de funções, que são o gargalo de muitas técnicas mais complexas da área.

- [todo: resta melhorar muito ainda esta aula, porém o esboço é este]

- [todo: HZ identities$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \cdots$]

Aula 14

To Do

1. Explicar melhor como retas e planos sao nao-lineares
2. Figura de transformações nao-lineares e nao-diferenciaveis (rasgar, formar bicos, etc)
3. enviar para aluno carlos
4. Aula: Transformacoes rigidas: Rotacao e translacao
5. Aula: Coordenadas homogeneas
6. Aula: SVD: teorema e demonstracao, algumas figs, relembrar autovalores
7. Aula: Aplicacao em imagens
8. Aula: Aplicacao em cameras
9. Aula: Aplicacao em reconhecimento