

Walisson Pereira Lorigiola

# Compressão de Imagens Digitais via DVS

FLORIANÓPOLIS - SC

Dezembro de 2004

Walisson Pereira Lorigiola

## Compressão de Imagens Digitais via DVS


Trabalho de Conclusão de Curso apresentado ao  
Curso de Matemática - Habilitação Licenciatura  
Departamento de Matemática  
Centro de Ciências Físicas e Matemáticas  
Universidade Federal de Santa Catarina

Orientador: Lício Hernanes Bezerra, Dr.

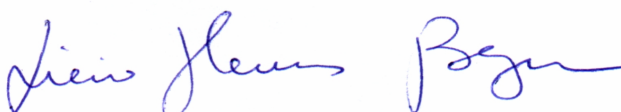
FLORIANÓPOLIS - SC


Dezembro de 2004

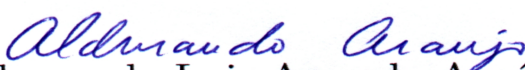
Esta Monografia foi julgada adequada como **TRABALHO DE CONCLUSÃO DE CURSO** no Curso de Matemática - Habilitação Licenciatura e aprovado em sua forma final pela Banca Examinadora designada pela Portaria nº 65/SCG/04.

  
Prof<sup>a</sup> Carmem Suzane Comitre Gimenez, Ms.  
Professora da disciplina

Banca examinadora

  
Prof. Lício Hernanes Bezerra, Dr.  
Professor Orientador

  
Prof. Daniel Norberto Kozakevich, Dr.  
Professor Membro

  
Prof. Aldrovando Luis Azeredo Araújo, Dr.  
Professor Membro

# *Dedicatória*

Com admiração dedico este trabalho  
aos meus pais NAIDE e JOÃO, por sempre  
acreditarem em mim;  
à minha irmã FORTUNATA e ao meu cunhado ELJI, pela ajuda  
e estímulos em todas as horas;  
à minha cunhada VERA e ao meu irmão JEFERSON que,  
apesar de tudo que a vida lhe impôs,  
mostra a cada dia que é uma pessoa capaz e vencedora;  
e, finalmente,  
aos meus sobrinhos VIVIANI, SUZI e PETERSON, que sempre  
me alegram com seus sorrisos despreocupados  
de criança; e a minha sobrinha ALINY.

# *Agradecimentos*

Agradeço a Deus;  
aos meus pais;  
à minha irmã e ao meu cunhado, pelo acolhimento;  
ao professor Lício, por orientar-me;  
ao amigo Luiz Camisão, por propiciar a minha inclusão digital;  
aos membros da banca professores Daniel e Aldrovando;  
aos professores do curso, em especial,  
à professora Carmem e ao professor Clóvis;  
aos amigos:  
Alexandra, Elen, Felipe, João, Karla, Louise, Priscila e Rui,  
pois estiveram ao meu lado nas diversas etapas deste curso;  
aos demais amigos e colegas;  
aos funcionários da Coordenadoria do Curso: Sílvia, Iara e Alcino.

# *Sumário*

## **Lista de Figuras**

<b>1</b>	<b>Introdução</b>	p. 9
<b>2</b>	<b>Produto de Matrizes</b>	p. 10
2.1	Matrizes blocadas (ou em bloco) . . . . .	p. 12
<b>3</b>	<b>Conhecendo melhor algumas normas</b>	p. 14
3.1	Norma de um Vetor . . . . .	p. 14
3.2	Norma de uma Matriz . . . . .	p. 16
<b>4</b>	<b>Decomposição QR</b>	p. 18
4.1	Processo de Gram-Schmidt . . . . .	p. 18
4.1.1	Processo de Gram-Schmidt modificado . . . . .	p. 19
4.2	Decomposição QR pelo Processo de Gram-Schmidt . . . . .	p. 21
4.3	Reflexões de Householder . . . . .	p. 23
4.4	Decomposição QR por Reflexões de Householder . . . . .	p. 27
<b>5</b>	<b>Problema de Quadrados Mínimos</b>	p. 28
<b>6</b>	<b>Dois Teoremas Fundamentais</b>	p. 31
6.1	O Teorema de Schur . . . . .	p. 31
6.2	O Teorema Espectral . . . . .	p. 32
<b>7</b>	<b>Decomposição em Valores Singulares</b>	p. 33

7.1 Os Valores Singulares de uma Matriz . . . . .	p. 33
7.2 A Decomposição em Valores Singulares - o Teorema . . . . .	p. 37
<b>8 Compressão de Imagem Usando a DVS</b>	p. 41
<b>9 Conclusão</b>	p. 50
<b>Referências Bibliográficas</b>	p. 51

# *Lista de Figuras*

1	Bola unitária nas normas 2, 1 e $\infty$ . . . . .	p. 15
2	Visualização geométrica do processo de Gram-Schmidt . . . . .	p. 19
3	Reflexão em relação à reta que contém $O$ e $B$ . . . . .	p. 23
4	Obtenção de $v$ por uma reflexão de $u$ em relação a $W^\perp$ . . . . .	p. 24
5	Visualização de transformação de $\mathbb{R}^3$ em $\mathbb{R}^2$ . . . . .	p. 34
6	Cores de $A$ no mapa <code>default</code> . . . . .	p. 43
7	Cores de $A$ no mapa <code>gray</code> . . . . .	p. 43
8	Uma imagem $20 \times 20$ . . . . .	p. 44
9	Compressão da figura 8 com $p = 2$ (esquerda) $p = 3$ (direita) . . . . .	p. 45
10	Compressão da figura 8 com $p = 4$ (esquerda) $p = 5$ (direita) . . . . .	p. 45
11	Uma imagem $576 \times 1280$ . . . . .	p. 46
12	Compressão da figura 11 com $p = 10$ (esquerda) e com $p = 20$ (direita) . . . . .	p. 46
13	Imagem original (esquerda) e sua compressão com $p = 40$ (direita) . . . . .	p. 47
14	Foto do Departamento em <code>jpg</code> (esquerda) e no formato <code>mat</code> (direita) . . . . .	p. 48
15	Compressão da figura 14 com $p = 10$ (esquerda) e com $p = 50$ (direita) . . . . .	p. 49
16	Imagem original (esquerda) e sua compressão com $p = 100$ (direita) . . . . .	p. 49



# 1 *Introdução*

Com a evolução dos meios de comunicação digital a quantidade de dados exigidos para representar, sem perda de informação, uma imagem inviabilizaria utilizações comuns, como uma foto de amigos exposta na Internet.

Armazenar ou transmitir grandes quantidades de dados usando recursos limitados de processamento e gerar imagens realmente relevantes com perda de informação intencional são grandes desafios. Problemas como este motivam pesquisadores a estudar diversas maneiras de contorná-los; tornam-se necessários métodos eficazes para o armazenamento, o processamento e a recuperação de imagens. Tais métodos devem ser aplicados a qualquer imagem e, para isso, esses métodos de redução da redundância de dados são estudados.

Neste trabalho iremos explorar a chamada decomposição em valores singulares (DVS). Uma imagem digital é representada por uma matriz. Diminuindo-se a redundância de dados dessa matriz, diminuimos a redundância de dados necessários à armazenagem ou transmissão de uma imagem. Tal efeito de eliminação de redundância é chamado de compressão.

Com este propósito, neste trabalho, apresentaremos alguns resultados básicos da álgebra linear computacional, para que possamos tratar da DVS, propriamente dita.

O foco desse trabalho não é, exatamente, o de examinar a eficiência da decomposição em valores singulares, aplicada à compressão de imagens comparada a outros métodos. Propomo-nos difundir a idéia que as decomposições que se estudam em Álgebra Linear servem para extrair das matrizes propriedades primordiais e, no caso da DVS, propriedades que traduzam quase fielmente uma imagem.

Assim, os capítulos 2, 3, 4, 5, 6 e 7, apresentam resultados de Álgebra Linear em uma seqüência que converge à DVS. No capítulo 8, apresentamos resultados de experimentos com compressão de imagens via DVS, nos que usamos MatLab.

## 2 Produto de Matrizes

A definição de produto usual de matrizes presta-se à descrição de sistemas lineares, como uma generalização de uma equação escalar do tipo

$$ax = b \quad (2.1)$$

em que  $a$ ,  $b$  são escalares,  $x$  é a solução escalar desejada. Ou seja, dado um sistema linear  $m \times n$

$$\begin{array}{cccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m \end{array}$$

é desejável escrever o sistema em uma forma semelhante à equação 2.1, isto é, como uma equação matricial

$$Ax = b$$

onde  $A = (a_{ij})$  é conhecida,  $x$  é uma matriz  $n \times 1$  de incógnitas e  $b$  é uma matriz  $m \times 1$  que representa os elementos à direita dos sinais de igualdade. Definimos, então,

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

e

$$Ax = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix}$$

O  $i$ -ésimo elemento de  $Ax$  é determinado pela  $i$ -ésima linha de  $A$ . Os elementos daquela linha são multiplicados pelos elementos correspondentes em  $x$  e os  $n$  produtos são, então,

somados. Ou seja, por esta “regra”, o  $i$ -ésimo elemento de  $Ax$  é determinado por:

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n$$

Pela familiaridade com produtos internos reconhecemos essa fórmula, simplesmente, como o produto interno da  $n$ -upla correspondente a  $i$ -ésima linha de  $A$  com a  $n$ -upla correspondente ao vetor  $x$

$$\begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n.$$

Assim, é possível definir um produto de uma matriz  $A$  por uma outra matriz  $B$  se o número de colunas de  $A$  é igual ao número de linhas de  $B$ : cada coluna  $B(:, k)$  dá origem a  $AB(:, k)$  que é o resultado do produto da matriz  $A$  pela matriz coluna  $B(:, j)$ . Ou melhor, dadas as matrizes  $A = (a_{ij})_{m \times n}$  e  $B = (b_{ij})_{n \times r}$ , o produto  $AB = C$  é definido do seguinte modo: cada coluna de  $C$  é o produto de  $A$  pela coluna de  $B$  correspondente  $C(:, j) = AB(:, j)$ . Desse modo, cada entrada  $(c_{ij})$  da matriz  $C$  é definida por:

$$c_{ij} = \sum_{k=1}^n a_{ik}b_{kj}$$

### Observações:

1. Observe que  $C(i, :)$  a  $i$ -ésima linha de  $C$ , é dada por

$$C(i, :) = A(i, :)B$$

2. Uma forma não usual de representar o produto de matrizes acima é a seguinte:

$$C = \sum_{i=1}^p A(:, i)B(i, :)$$

## 2.1 Matrizes blocadas (ou em bloco)

Matrizes em bloco podem ser multiplicadas de modo análogo, como se os blocos fossem entradas numéricas, desde que, para um produto  $AB$ , a partição das colunas de  $A$  combine com a partição das linhas de  $B$ . Veja o exemplo que se segue.

- **Exemplo:** Sejam

$$A = \begin{pmatrix} 2 & -3 & 1 & 0 & -4 \\ 1 & 5 & -2 & 3 & -1 \\ 0 & -4 & -2 & 7 & -1 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 6 & 4 \\ -2 & 1 \\ -3 & 7 \\ -1 & 3 \\ 5 & 2 \end{pmatrix}$$

blocadas da seguinte maneira:

- blocos de  $A$ :  $A_{11} = \begin{pmatrix} 2 & -3 & 1 \\ 1 & 5 & -2 \end{pmatrix}$ ,  $A_{12} = \begin{pmatrix} 0 & -4 \\ 3 & -1 \end{pmatrix}$ ,  $A_{21} = \begin{pmatrix} 0 & -4 & -2 \end{pmatrix}$  e  $A_{22} = \begin{pmatrix} 7 & -1 \end{pmatrix}$ ;

- blocos de  $B$ :  $B_1 = \begin{pmatrix} 6 & 4 \\ -2 & 1 \\ -3 & 7 \end{pmatrix}$  e  $B_2 = \begin{pmatrix} -1 & 3 \\ 5 & 2 \end{pmatrix}$

O produto  $AB$  pode ser escrito como

$$AB = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} A_{11}B_1 + A_{12}B_2 \\ A_{21}B_1 + A_{22}B_2 \end{pmatrix}$$

Agora,

$$A_{11}B_1 = \begin{pmatrix} 15 & 12 \\ 2 & -5 \end{pmatrix}, \quad A_{12}B_2 = \begin{pmatrix} -20 & -8 \\ -8 & 7 \end{pmatrix},$$

$$A_{21}B_1 = \begin{pmatrix} 14 & -18 \end{pmatrix} \quad \text{e} \quad A_{22}B_2 = \begin{pmatrix} -12 & 19 \end{pmatrix}.$$

Logo, o bloco superior de  $AB$  é  $\begin{pmatrix} -5 & 4 \\ -6 & 2 \end{pmatrix}$  e o seu bloco inferior é  $\begin{pmatrix} 2 & 1 \end{pmatrix}$ .

$$\text{Portanto } AB = \begin{pmatrix} -5 & 4 \\ -6 & 2 \\ 2 & 1 \end{pmatrix}$$

O produto de matrizes em blocos é a forma mais genérica de considerar o produto de duas matrizes. Este produto terá importância ímpar no capítulo 7.2 onde faremos considerações sobre a DVS reduzida.

## 3 *Conhecendo melhor algumas normas*

### 3.1 Norma de um Vetor

Uma norma no espaço  $\mathbb{R}^n$  é uma função  $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$  de modo que sejam válidas as seguintes propriedades:

- $\|v\| \geq 0$ , para todo  $v \in \mathbb{R}^n$ ;
- $\|v\| = 0 \iff v = 0$ ; (se não for válida esta propriedade, diz-se então que é uma semi-norma);
- $\|\alpha v\| = |\alpha| \|v\|$ , para todo  $\alpha \in \mathbb{R}$ , e para todo  $v \in \mathbb{R}^n$ ;
- $\|v + u\| \leq \|v\| + \|u\|$ , (desigualdade triangular)

Dessa desigualdade, segue que  $\|v - u\| \geq \left| \|v\| - \|u\| \right|$ .

Exemplos de normas são os seguintes:

- $\|v\|_2 = \sqrt{\langle v, v \rangle} = \sqrt{v^T v} = \sqrt{\sum_{i=1}^n v_i^2}$ ; (norma euclidiana)
- $\|v\|_\infty = \max_{i=1, \dots, n} |v_i|$ ; (norma do máximo)
- $\|v\|_1 = \sum_{i=1}^n |v_i|$

Por exemplo: Se  $v = \begin{pmatrix} 1 & -1 & 2 \end{pmatrix}$ ,  $\|v\|_2 = \sqrt{6}$ ,  $\|v\|_\infty = 2$  e  $\|v\|_1 = 4$ .

**Definição 3.1.1.** Duas normas  $\|\cdot\|_\alpha$  e  $\|\cdot\|_\beta$  são equivalentes se existirem constantes  $a$  e  $b$  tais que para qualquer  $v$ :

$$\|v\|_\alpha \leq a\|v\|_\beta$$

$$\|v\|_\beta \leq b\|v\|_\alpha$$

Num espaço de dimensão finita, como é o caso do  $\mathbb{R}^n$  todas as normas são equivalentes, ou seja, se  $\|\cdot\|_\alpha$  e  $\|\cdot\|_\beta$  são normas do  $\mathbb{R}^n$ , então existe constantes positivas  $c_1$  e  $c_2$  tais que  $c_1\|v\|_\alpha \leq \|v\|_\beta \leq c_2\|v\|_\alpha$ .

**Definição 3.1.2.** Uma seqüência  $(v^{(k)})$  de vetores do  $\mathbb{R}^n$  converge a um vetor  $v$ , quando  $k \rightarrow \infty$ , se existe uma norma  $\|\cdot\|$  tal que o  $\lim_{k \rightarrow \infty} \|v^{(k)} - v\| = 0$ .

Esta definição está bem posta porque todas as normas são equivalentes. Além disso,

$$\lim_{k \rightarrow \infty} v^{(k)} = v \Leftrightarrow \lim_{k \rightarrow \infty} v_i^{(k)} = v_i \quad \text{para todo } 1 \leq i \leq n$$

$$\Leftrightarrow \lim_{k \rightarrow \infty} \|v^{(k)} - v\| = 0$$

**Exemplos:**

1. Considere para todo  $k \in \mathbb{Z}^+$ ,  $v^{(k)} = \left(\frac{1}{k} \quad 1 \quad \frac{1}{k^2}\right)$ , assim  $\lim_{k \rightarrow \infty} v^{(k)} = \left(0 \quad 1 \quad 0\right)$
2. Considere a bola fechada  $B = \{v \in \mathbb{R}^2, \text{ tal que } \|v\| = 1\}$ . A representação de  $B$  em  $\mathbb{R}^2$  com respeito as normas 2, 1 e  $\infty$  são as seguintes:

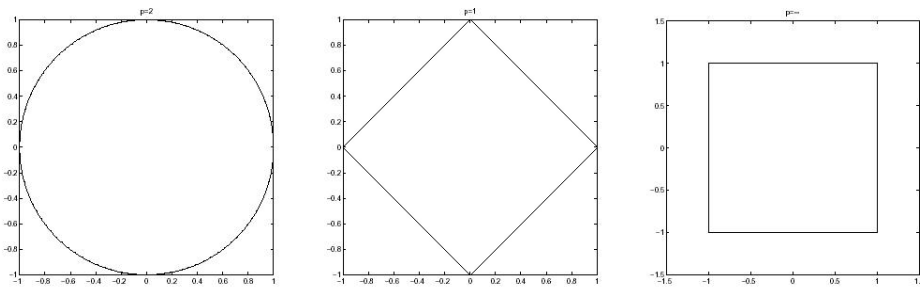


Figura 1: Bola unitária nas normas 2, 1 e  $\infty$

## 3.2 Norma de uma Matriz

Podemos pensar uma matriz  $m \times n$  como um vetor de  $m \times n$  componentes (ordenando os elementos por linha ou por coluna), ou seja, uma norma matricial é uma função  $\|\cdot\| : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ , que deve satisfazer:

- $\|A\| \geq 0$ , para todo  $A \in \mathbb{R}^{m \times n}$ ;
- $\|A\| = 0 \iff A = 0$ ;
- $\|\alpha A\| = |\alpha| \|A\|$ , para todo  $\alpha \in \mathbb{R}$ , e para todo  $v \in \mathbb{R}^{m \times n}$ ;
- $\|A + B\| \leq \|A\| + \|B\|$ .

Mas, acrescentamos uma nova propriedade para que uma norma seja norma de matrizes:

- $\|AB\| \leq \|A\| \|B\|$

Dizemos que a norma de matrizes  $\|\cdot\|_M$  é compatível com a norma de vetores  $\|\cdot\|$  se  $\forall A \in \mathbb{R}^{m \times n}$  e  $\forall v \in \mathbb{R}^n$

$$\|Av\| \leq \|A\|_M \|v\|$$

### Exemplos

- Norma matricial compatível com a norma vetorial  $\infty$

$$\|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|;$$

pois  $\|Av\|_\infty \leq \|A\|_\infty \|v\|_\infty$

- Norma matricial compatível com a norma vetorial 1

$$\|A\|_1 = \max_j \sum_{i=1}^m |a_{ij}|;$$

- Norma matricial compatível com a norma euclidiana (norma-2)<sup>1</sup>, chamada de norma espectral

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}$$

em que  $\lambda_{\max}(A^T A)$  indica o máximo autovalor de  $A^T A$ ;

---

<sup>1</sup>Veremos no capítulo 7.1 que a norma-2 é, tão somente, o maior valor singular de  $A$ .



- Uma outra norma matricial compatível com a norma euclidiana é a norma de Frobenius

$$\|A\|_F = \sqrt{\sum_i \sum_j (a_{ij})^2} = \sqrt{\text{tr}(A^T A)}$$

em que  $\text{tr}(A^T A)$  é a soma dos elementos da diagonal de  $A^T A$ , ou seja, o traço de  $A^T A$ .

**Exemplo:** Seja  $A = \begin{pmatrix} 1 & 1 & -1 \\ 0 & -3 & 2 \end{pmatrix}$ . Logo:  $\|A\|_F = 4$ ,  $\|A\|_1 = 4$ ,  $\|A\|_\infty = 5$  e  $\|A\|_2 = \sqrt{8 + 5\sqrt{2}}$

**Definição 3.2.1.** Dada uma seqüência de matrizes  $\{A_k\}$  dizemos que ela converge a um matriz  $A$ , quando  $k \rightarrow \infty$ , se existe uma norma matricial  $\|\cdot\|$  tal que  $\lim_{k \rightarrow \infty} \|A_k - A\| = 0$ .

Tal definição é abrangente, pois todas as normas matriciais são equivalentes. Além disso

$$\lim_{k \rightarrow \infty} A_k = A \iff \forall_{i,j} \lim_{k \rightarrow \infty} a_{ij}^{(k)} = a_{ij}$$

## 4 Decomposição $QR$

Neste capítulo apresentaremos a chamada decomposição  $QR$  de uma matriz  $A_{m \times n}$ , isto é, um produto de duas outras matrizes  $Q$  e  $R$ , em que  $Q$  é uma matriz ortogonal e  $R$  é uma matriz triangular superior.

Para tal decomposição, faremos uso de dois métodos: o processo de ortogonalização de Gram-Schmidt e o método de ortogonalização via reflexões de Householder. Descreveremos, agora, cada um dos processos e a forma de obtenção da decomposição  $QR$ , via cada método.

### 4.1 Processo de Gram-Schmidt

A partir de uma base para um espaço vetorial qualquer (com produto interno  $\langle, \rangle$ ), o processo de Gram-Schmidt é usado para gerar uma base ortogonal desse espaço (em relação a esse produto interno). Inicialmente descreveremos esse processo para uma base  $\beta = \{v_1, v_2\}$ .

Seja  $u_1 = v_1$ . Teremos de encontrar a partir de  $v_2$  um novo vetor  $u_2$  ortogonal a  $u_1$ , isto é,  $\langle u_1, u_2 \rangle = 0$ . Para tal tomamos  $u_2 = v_2 - \alpha_{12}u_1$ , onde  $\alpha_{12}$  é um número escolhido de modo que  $\langle u_1, u_2 \rangle = 0$ , ou melhor,  $\langle u_1, v_2 - \alpha_{12}u_1 \rangle = 0$  o que resulta em  $\alpha_{12} = \frac{\langle v_2, u_1 \rangle}{\langle u_1, u_1 \rangle}$ . ficamos então com:

$$u_1 = v_1$$

$$u_2 = v_2 - \alpha_{12}u_1 = v_2 - \frac{\langle u_1, v_2 \rangle}{\langle u_1, u_1 \rangle} u_1$$

Observemos que  $u_2$  foi obtido de  $v_2$ , subtraindo-se deste a projeção ortogonal de  $v_2$  na direção de  $u_1$ ,  $\frac{\langle u_1, v_2 \rangle}{\langle u_1, u_1 \rangle} u_1$

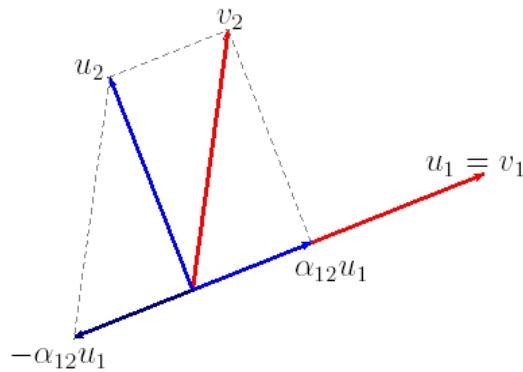


Figura 2: Visualização geométrica do processo de Gram-Schmidt

Como definido para uma base  $\beta$  com dois vetores, podemos estender o processo para uma base com um número arbitrário de vetores. Vamos, agora, transpor esse processo para os casos em que as bases são formadas por vetores do  $\mathbb{R}^n$  e o produto interno é o usual.

Sejam  $\{v_1, \dots, v_n\}$  vetores colunas linearmente independentes de uma matriz  $A \in \mathbb{R}^{m \times n}$ .

Então

$$u_1 = v_1$$

$$u_2 = v_2 - \alpha_{12}u_1$$

$$u_3 = v_3 - \alpha_{13}u_1 - \alpha_{23}u_2$$

$\vdots$

$$u_n = v_n - \alpha_{1n}u_1 - \alpha_{2n}u_2 - \dots - \alpha_{(n-1)n}u_{(n-1)}$$

Genericamente, definindo  $\alpha_{ij} = \frac{u_i^T v_j}{u_i^T u_i}$ , obtemos:

$$u_j = v_j - \sum_{i=1}^{(j-1)} \alpha_{ij}u_i$$

#### 4.1.1 Processo de Gram-Schmidt modificado

O processo de Gram-Schmidt não produz, em geral, resultados precisos quando trabalhamos com aritmética de precisão finita. Na prática, os vetores resultantes não são perpendiculares devido a erros de aproximação no cálculo de  $u_1, u_2, u_3, \dots, u_n$ . Podemos obter melhor precisão numérica usando uma versão modificada do método de Gram-

Schmidt, em que:

- na primeira etapa, chamando  $u_1 = v_1$ , subtraímos de cada vetor  $v_i$  sua projeção ortogonal sobre  $u_1$ . Temos, então, que:

$$v_2^{(1)} = v_2 - \alpha_{12}u_1$$

$$v_3^{(1)} = v_3 - \alpha_{13}u_1$$

$$v_4^{(1)} = v_4 - \alpha_{14}u_1$$

⋮

$$v_n^{(1)} = v_n - \alpha_{1n}u_1$$

em que  $\alpha_{1j} = \langle u_1, v_j \rangle$ . Definimos  $u_2 = v_2^{(1)}$

- Na segunda etapa, modificamos os  $(n - 2)$  vetores restantes de modo a torná-los ortogonais a  $u_2$  (observe que pela primeira etapa, eles já são ortogonais a  $u_1$ ).

$$v_3^{(2)} = v_3^{(1)} - \alpha_{23(1)}u_2$$

$$v_4^{(2)} = v_4^{(1)} - \alpha_{24(1)}u_2$$

$$v_5^{(2)} = v_5^{(1)} - \alpha_{25(1)}u_2$$

⋮

$$v_n^{(2)} = v_n^{(1)} - \alpha_{2n(1)}u_2$$

em que  $\alpha_{2j} = \langle u_2, v_j \rangle$ ,  $j \geq 3$ . Definimos  $u_3 = v_3^{(2)}$ .

Analogamente, determinamos  $u_4, u_5, \dots, u_n$ .

O processo de Gram-Schmidt modificado pode, também, ser descrito por meio de matrizes de projeção:

$$v_1^{(0)} = u_1$$

$$v_2^{(1)} = v_2^{(0)} - \frac{u_1^T v_2}{u_1^T u_1} u_1 = v_2^{(0)} - \frac{(v_1^{(0)})^T v_2^{(0)}}{(v_1^{(0)})^T v_1^{(0)}} v_1^{(0)} = v_2^{(0)} - P_1 v_2^{(0)} = (I - P_1) v_2^{(0)}$$

onde  $P = v_1^{(0)}(v_1^{(0)})^T$ .

Desta forma, para um  $k$  qualquer, temos:

$$v_k^{(0)} = v_k$$

$$v_k^{(1)} = (I - P_1)v_k^{(0)}$$

$$v_k^{(2)} = (I - P_2)v_k^{(1)} = (I - P_2)(I - P_1)v_k^{(0)}$$

⋮

$$v_k^{(n)} = (I - P_n)v_k^{(n-1)} = (I - P_n)(I - P_{n-1}) \dots (I - P_2)(I - P_1)v_k^{(0)}$$

E  $u_k$  será tomado da mesma forma que anteriormente, ou seja,  $u_k = v_k^{(k-1)}$ .

Se o processo de Gram-Schmidt (modificado ou não) é aplicado aos vetores colunas de uma matriz  $A_{m \times p}$  de posto  $p$ , então podemos obter uma fatoração  $QR$  de  $A$ .

## 4.2 Decomposição QR pelo Processo de Gram-Schmidt

A decomposição QR é uma interpretação matricial do processo de Gram-Schmidt. Suponha que  $v_1, v_2, \dots, v_p$  são as  $p$  colunas de uma matriz  $A$  e considere a tradicional implementação do processo de Gram-Schmidt. Temos pela definição do processo que cada vetor  $v_j$  pode ser escrito como:

$$v_j = u_1\alpha_{1,j} + u_2\alpha_{2,j} + \dots + u_{j-1}\alpha_{j-1,j} + u_j$$

O que pode ser reescrito, em notação matricial, por um produto de matrizes:

$$\begin{pmatrix} \vdots & & \vdots \\ v_1 & \dots & v_p \\ \vdots & & \vdots \end{pmatrix} = \begin{pmatrix} \vdots & & \vdots \\ u_1 & \dots & u_p \\ \vdots & & \vdots \end{pmatrix} \begin{pmatrix} 1 & \alpha_{1,2} & \alpha_{1,3} & \dots & \alpha_{1,p} \\ 0 & 1 & \alpha_{2,3} & \dots & \alpha_{2,p} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Se dividirmos cada vetor  $u_i$  por sua norma, essa igualdade fica:

$$\begin{pmatrix} \vdots & & \vdots \\ v_1 & \dots & v_p \\ \vdots & & \vdots \end{pmatrix} = \underbrace{\begin{pmatrix} \vdots & & \vdots \\ \frac{u_1}{\|u_1\|} & \dots & \frac{u_p}{\|u_p\|} \\ \vdots & & \vdots \end{pmatrix}}_Q \underbrace{\begin{pmatrix} \|u_1\| & & 0 \\ & \ddots & \\ 0 & & \|u_p\| \end{pmatrix}}_R \begin{pmatrix} 1 & \alpha_{1,2} & \dots & \alpha_{1,p} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

A matriz  $\begin{pmatrix} \vdots & & \vdots \\ \frac{u_1}{\|u_1\|} & \dots & \frac{u_p}{\|u_p\|} \\ \vdots & & \vdots \end{pmatrix}$ , que denotaremos por  $Q$ , tem colunas ortonormais (ou seja,

$Q^T Q = I$ ). O produto  $\begin{pmatrix} \|u_1\| & & 0 \\ & \ddots & \\ 0 & & \|u_p\| \end{pmatrix} \begin{pmatrix} 1 & \alpha_{1,2} & \dots & \alpha_{1,p} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$  é uma matriz triangular superior,

que denotaremos por  $R$ . Assim  $A = QR$ .

**Teorema 4.2.1 (Teorema da Decomposição QR por Gram-Schmidt).** *Seja  $A$  uma matriz  $m \times p$  de posto  $p$ . Então  $A$  pode ser escrita como uma decomposição QR,  $A = QR$ , em que:*

- $Q$  é uma matriz  $m \times p$  e suas colunas são ortonormais;
- $R$  é uma matriz  $p \times p$ , triangular-superior, e de posto  $p$ .

**Exemplo:** Considere a matriz  $A_{4 \times 4}$  da qual as 4 colunas originais são vetores  $v_1, v_2, v_3, v_4$ .

$$A = \begin{pmatrix} 1 & 2 & 0 & -1 \\ 1 & -1 & 3 & 2 \\ 1 & -1 & 3 & 2 \\ -1 & 1 & -3 & 1 \end{pmatrix}$$

Aplicando a  $v_1, v_2, v_3, v_4$  o processo de Gram-Schmidt temos:

$$u_1 = v_1 = \begin{pmatrix} 1 & 1 & 1 & -1 \end{pmatrix}^T$$

$$u_2 = v_2 - \alpha_{12}u_1 = v_2 - \frac{\langle u_1, v_2 \rangle}{\langle u_1, u_1 \rangle} u_1 = \begin{pmatrix} \frac{9}{4} & -\frac{3}{4} & -\frac{3}{4} & \frac{3}{4} \end{pmatrix}^T$$

$$u_3 = v_3 - \alpha_{13}u_1 - \alpha_{23}u_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}^T$$

$$u_4 = v_4 - \alpha_{14}u_1 - \alpha_{24}u_2 - \alpha_{34}u_3 = \begin{pmatrix} 0 & 1 & 1 & 2 \end{pmatrix}^T$$

Portanto,

$$Q = \begin{pmatrix} \frac{u_1}{\|u_1\|} & \frac{u_2}{\|u_2\|} & \frac{u_4}{\|u_4\|} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{1}{2} & -\frac{\sqrt{3}}{6} & \frac{\sqrt{6}}{6} \\ \frac{1}{2} & -\frac{\sqrt{3}}{6} & \frac{\sqrt{6}}{6} \\ -\frac{1}{2} & \frac{\sqrt{3}}{6} & \frac{\sqrt{6}}{3} \end{pmatrix}$$

$$R = \begin{pmatrix} 2 & -\frac{1}{2} & \frac{9}{2} & 1 \\ 0 & \frac{3\sqrt{3}}{2} & -\frac{3\sqrt{3}}{2} & -\sqrt{3} \\ 0 & 0 & 0 & \sqrt{6} \end{pmatrix}$$

### 4.3 Reflexões de Householder

Seja  $P$  uma matriz ortogonal (isto é,  $P \in \mathbb{R}^{n \times n}$  e  $P^T P = I$ ). Sabemos que as seguintes propriedades são equivalentes a  $P$  ser ortogonal.

- $\langle Pu, Pv \rangle = \langle u, v \rangle$  para todo  $u$  e  $v$ ; assim o ângulo entre  $Pu$  e  $Pv$  é igual aquele entre  $u$  e  $v$ .
- $\|Pu\|_2 = \|u\|_2$  para todo  $u$ ; assim o comprimento de  $Pu$  é igual ao comprimento de  $u$ .

Portanto,  $P$  preserva comprimento e ângulo.

Exemplos de matrizes ortogonais são as matrizes correspondentes às operações de reflexão em relação a um subespaço de  $\mathbb{R}^n$ .

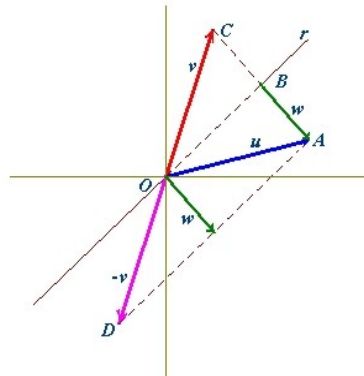


Figura 3: Reflexão em relação à reta que contém  $O$  e  $B$

Suponha que desejamos refletir um vetor real  $u$  sobre a reta  $r$  que passa pela origem. Temos duas soluções:

- uma que resulta no vetor  $\overrightarrow{OC}$ ;
- outra, que resulta no vetor  $\overrightarrow{OD}$ .

Observe que  $\overrightarrow{OC} = -\overrightarrow{OD}$  e que  $\|\overrightarrow{OC}\| = \|\overrightarrow{OD}\| = \|u\|$ .

Também temos que  $\overrightarrow{OA} = \overrightarrow{OC} + 2\overrightarrow{BA}$ , em que  $\overrightarrow{BA}$  é a projeção ortogonal de  $u$  sobre  $\overrightarrow{CA}$ , isto é,

$$\overrightarrow{BA} = \frac{\langle \overrightarrow{OA}, \overrightarrow{CA} \rangle}{\|\overrightarrow{CA}\|^2} \overrightarrow{CA}.$$

Denotando  $\overrightarrow{CA} = u - v$  por  $w$ , temos que  $u = v + 2P_w u$ , em que  $P_w u = \overrightarrow{BA} = \frac{u, w}{\|w\|^2} w$ , ou  $\frac{w w^T u}{w^T w}$ , em notação matricial, em que  $u$  e  $w$  são matrizes  $n \times 1$ .

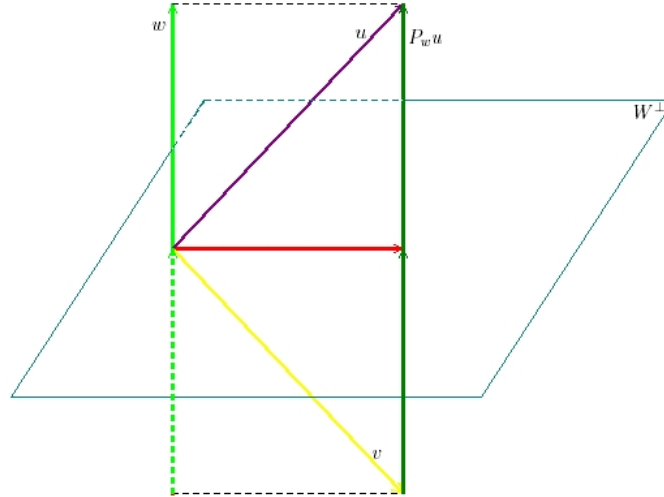


Figura 4: Obtenção de  $v$  por uma reflexão de  $u$  em relação a  $W^\perp$

Ou seja,

$$v = \left( I - 2 \frac{w w^T}{w^T w} \right) u$$

Portanto,

$$P_w = \left( I - 2 \frac{w w^T}{w^T w} \right).$$

Observe que  $P_w$  é a matriz associada ao operador que reflete um vetor  $u$  em relação a  $[w]^\perp$ , complemento ortogonal de  $[w]$ . Essa reflexão é dita reflexão de Householder.

### Exemplos

- Considere o vetor  $u = \begin{pmatrix} 3 & 4 \end{pmatrix}^T$  e o problema de transformá-lo, por uma reflexão de Householder, em um múltiplo do vetor  $e_1$ .

Resolução:

Para solucionar tal problema, definiremos o vetor  $w$  em termos do vetor  $u$  e de um vetor da reta em relação à qual  $u$  será refletido.

Como  $\|u\| = 5$ ,  $v$  pode ser  $\pm 5e_1$ . Vamos definir  $v = 5e_1$ . Logo,

$$w = u - 5e_1 = \begin{pmatrix} -2 & 4 \end{pmatrix}^T$$



$$H = I - 2 \frac{ww^T}{\|w\|^2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{2}{20} \begin{pmatrix} 4 & -8 \\ -8 & 16 \end{pmatrix} = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{pmatrix}$$

o que nos dá

$$v = Hu = \begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{pmatrix} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \end{pmatrix} = 5e_1$$

- Dada a matriz  $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ ,

(a) encontrar  $H_1$  tal que  $H_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$  seja múltiplo de  $e_1$ , e determinar  $A_1 = H_1 A$ ;

(b) determinar  $H_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \tilde{H}_2 & \\ 0 & & \end{pmatrix}$  tal que  $\tilde{H}_2 \begin{pmatrix} * \\ * \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ , em que  $\begin{pmatrix} * \\ * \end{pmatrix}$  é a

submatriz  $A_1$ :  $A_1(2 : 3, 2)$ .

Resolução:

- Para solucionar o item (a) procederemos de mesma forma que no exemplo anterior.

Temos

$$u_1 = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix}^T$$

$$e_1 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}^T$$

$$w_1 = u_1 - \|u_1\|e_1 = \left( (1 - \sqrt{2}) \quad 1 \quad 0 \right)^T$$

Como  $H_1 = I - \frac{2}{\|w_1\|^2} w_1 w_1^T$

$$H_1 = \begin{pmatrix} \frac{-2-\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- Determinaremos, agora,  $A_1$  para solucionar o item (b)

$$A_1 = H_1 A = \begin{pmatrix} \frac{-2-\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{-\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & \frac{-2-\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{2}}{2} & \frac{-\sqrt{2}}{2} \\ 0 & 1 & 1 \end{pmatrix}$$

Tendo  $A_1$  podemos encontrar  $\tilde{H}_2$  de modo a transformar  $u_2 = \left(\frac{\sqrt{2}}{2} \ 1\right)^T$  em múltiplo de  $e_1 = \begin{pmatrix} 1 & 0 \end{pmatrix}^T$ . Procedendo como no item (a),

$$w_2 = u_2 - \|u_2\|e_1 = \left(\frac{\sqrt{2}-\sqrt{6}}{2} \ 1\right)^T$$

$$\tilde{H}_2 = \begin{pmatrix} \frac{\sqrt{12}}{6} & \frac{\sqrt{6}}{3} \\ \frac{\sqrt{6}}{3} & \frac{-\sqrt{12}}{6} \end{pmatrix}$$

$$\text{Logo, } H_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{12}}{6} & \frac{\sqrt{6}}{3} \\ 0 & \frac{\sqrt{6}}{3} & \frac{-\sqrt{12}}{6} \end{pmatrix}$$

## 4.4 Decomposição QR por Reflexões de Householder

A decomposição  $QR$  por matrizes de Householder é, consideravelmente, em termos computacionais, mais precisa do que a decomposição obtida pelo processo de Gram-Schmidt (conforme (2)).

Suponhamos  $A_{m \times n}$  uma matriz. Definiremos uma seqüência  $H_1, H_2, \dots, H_{n-1}$  de matrizes de Householder, tal que

$$H_{n-1}H_{n-2} \dots H_1 A = R$$

em que  $R$  é triangular superior. Equivalentemente  $A = QR$ , conquanto  $Q = H_1 H_2 \dots H_{n-1}$ .

- **Exemplo** Usaremos as matrizes  $H_1$  e  $H_2$  do segundo exemplo da seção 4.3 para decompor a matriz  $A$ , do citado exemplo, em  $QR$ .

Resolução:

$$\text{Temos: } A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad H_1 = \begin{pmatrix} \frac{-2-\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{-\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{e} \quad H_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{12}}{6} & \frac{\sqrt{6}}{3} \\ 0 & \frac{\sqrt{6}}{3} & \frac{-\sqrt{12}}{6} \end{pmatrix}.$$

Efetuamos o produto  $H_1 H_2 A$  e obtemos a matriz  $R$

$$H_1 H_2 A = \begin{pmatrix} -1 & \frac{-2-\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{6}}{2} & \frac{\sqrt{6}}{6} \\ 0 & 0 & \frac{-2\sqrt{3}}{3} \end{pmatrix} = R.$$

Por conseguinte,  $A = H_1^{-1} H_2^{-1} R$ . Como  $H_i^{-1} = H_i$ , temos  $A = H_1 H_2 R$ , ou seja,

$$Q = H_1 H_2 = \begin{pmatrix} \frac{-2-\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{-\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{12}}{6} & \frac{\sqrt{6}}{3} \\ 0 & \frac{\sqrt{6}}{3} & \frac{-\sqrt{12}}{6} \end{pmatrix} = \begin{pmatrix} \frac{-2-\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{2}}{2} & \frac{-\sqrt{6}}{6} & \frac{-\sqrt{3}}{3} \\ 0 & \frac{\sqrt{6}}{3} & \frac{-\sqrt{3}}{3} \end{pmatrix}$$

## 5 *Problema de Quadrados Mínimos*

Consideremos sistemas envolvendo mais equações do que incógnitas ( $m > n$ ). Então dado um sistema  $m \times n$   $Ax = b$  com  $m > n$ , não podemos esperar, em geral, encontrar um vetor  $x \in \mathbb{R}^n$  para o qual  $Ax$  seja igual  $b$ . Em vez disso, vamos procurar um vetor  $x$  para o qual  $Ax$  está o mais próximo possível de  $b$ .

Seja uma matriz  $A_{m \times n}$  com  $m > n$  e o vetor  $b \in \mathbb{R}^m$ . Para cada  $x \in \mathbb{R}^n$  podemos formar o resíduo  $r(x) = b - Ax$ . A distância entre  $b$  e  $Ax$  é dada por

$$\|b - Ax\| = \|r(x)\|$$

Queremos encontrar um vetor  $x$  para o qual  $\|r(x)\|$  é mínimo. Minimizar  $\|r(x)\|$  é equivalente a minimizar  $\|r(x)\|^2$ . O vetor  $\tilde{x}$  que nos proporciona tal minimização é uma solução de quadrados mínimos para o sistema  $Ax = b$ .

Se  $\tilde{x}$  é uma solução de quadrados mínimos para o sistema  $Ax = b$  e  $v = A\tilde{x}$ , então  $v$  é o vetor no espaço coluna de  $A$  mais próximo de  $b$ .

A ortogonalidade tem um papel importante na solução de quadrados mínimos. Se  $A_{m \times n}$  é uma matriz de posto  $n$ , então o problema de quadrados mínimos para  $Ax = b$  tem uma única solução  $\tilde{x}$  que é obtida resolvendo-se as equações normais  $A^T Ax = A^T b$ . A projeção ortogonal de  $v = A\tilde{x}$  é o vetor pertencente à imagem  $A$  mais próximo de  $b$ . O problema é facilmente solúvel quando os vetores colunas de  $A$  são ortonormais, pois  $A^T A = I$ . A solução dos quadrados mínimos é, nesse caso,  $\tilde{x} = A^T b$ .

Caso  $A$  não seja uma matriz com colunas ortonormais, obtemos a decomposição de  $A$  em um produto  $QR$ , onde  $Q$  é ortogonal e  $R$  é triangular superior.

**Teorema 5.0.1.** *Se  $A_{m \times n}$  é uma matriz de posto  $n$ , então a solução para o problema de quadrados mínimos  $Ax = b$  é dada por  $\tilde{x} = R^{-1}Q^T b$ , em que  $Q$  é ortogonal e  $R$  é triangular superior. A solução  $\tilde{x}$  pode ser obtida resolvendo-se o sistema  $Rx = Q^T b$  por*

*substituição.*

**Prova:** Seja  $\tilde{x}$  a solução do problema de quadrados mínimos para  $Ax = b$ . Então  $\tilde{x}$  satisfaz as equações normais

$$A^T Ax = A^T b$$

Decompondo  $A$  em um produto  $QR$ , essas equações ficam

$$(QR)^T(QR)x = (QR)^T b$$

ou seja

$$R^T(Q^T Q)Rx = R^T Q^T b$$

como  $Q$  é ortogonal, temos

$$R^T Rx = R^T Q^T b$$

$R$  é invertível, logo  $Rx = Q^T b$  e, portanto,

$$x = R^{-1}Q^T b$$

- **Exemplo** Encontre a solução de quadrados mínimos para o sistema

$$\begin{cases} x_1 - 2x_2 - x_3 = -1 \\ 2x_1 \quad \quad + x_3 = 1 \\ 2x_1 - 4x_2 + 2x_3 = 1 \\ 4x_1 \quad \quad \quad = -2 \end{cases} \quad (5.1)$$

**Resolução:** Reescrevendo o sistema em notação matricial

$$\begin{pmatrix} 1 & -2 & -1 \\ 2 & 0 & 1 \\ 2 & -4 & 2 \\ 4 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 1 \\ -2 \end{pmatrix}$$

Fazendo o uso do processo de Gram-Schmidt, decompondo em QR a matriz  $A$  dos coeficientes do sistema, ficamos com

$$A = \begin{pmatrix} 1 & -2 & -1 \\ 2 & 0 & 1 \\ 2 & -4 & 2 \\ 4 & 0 & 0 \end{pmatrix} = QR = \begin{pmatrix} \frac{1}{5} & \frac{-2}{5} & \frac{-4}{5} \\ \frac{2}{5} & \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{-4}{5} & \frac{2}{5} \\ \frac{4}{5} & \frac{2}{5} & \frac{-1}{5} \end{pmatrix} \begin{pmatrix} 5 & -2 & 1 \\ 0 & 4 & -1 \\ 0 & 0 & 2 \end{pmatrix}$$

Usando essa decomposição, obtemos

$$Q^T b = \begin{pmatrix} \frac{1}{5} & \frac{-2}{5} & \frac{-4}{5} \\ \frac{2}{5} & \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{-4}{5} & \frac{2}{5} \\ \frac{4}{5} & \frac{2}{5} & \frac{-1}{5} \end{pmatrix}^T \begin{pmatrix} -1 \\ 1 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix},$$

possibilitando-nos solucionar  $Rx = Q^T b$  facilmente com o uso de substituição:

$$\begin{pmatrix} 5 & -2 & 1 \\ 0 & 4 & -1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$$

Portanto,  $x = \begin{pmatrix} -\frac{2}{5} \\ 0 \\ 1 \end{pmatrix}$  é solução para o problema de quadrados mínimos, do sistema

5.1.

## 6 Dois Teoremas Fundamentais

### 6.1 O Teorema de Schur

**Teorema 6.1.1 (Teorema de Schur).** *Dada uma matriz  $A$ , existe uma matriz unitária  $Q$  e uma matriz triangular superior  $T$  tal que  $Q^H A Q = T$ , em que os autovalores de  $A$  são as entradas da diagonal de  $T$ .*

Para a prova deste teorema faremos o uso de indução. O teorema é obviamente válido se  $A$  é  $1 \times 1$ . Agora, sejam  $\lambda$  qualquer autovalor de  $A$  e  $u$  um correspondente autovetor normalizado ( $\|u\| = 1$ ). Escolhemos  $\tilde{U}$  de forma que  $U = \begin{pmatrix} u & \tilde{U} \end{pmatrix}$  seja uma matriz quadrada e unitária. Então:

$$U^H A U = \begin{pmatrix} u^H \\ \tilde{U}^H \end{pmatrix} A \begin{pmatrix} u & \tilde{U} \end{pmatrix} = \begin{pmatrix} u^H A \\ \tilde{U}^H A \end{pmatrix} \begin{pmatrix} u & \tilde{U} \end{pmatrix} = \begin{pmatrix} u^H A u & u^H A \tilde{U} \\ \tilde{U}^H A u & \tilde{U}^H A \tilde{U} \end{pmatrix}.$$

Agora, podemos reescrever

$$u^H A u = u^H \lambda u = \lambda u^H u = \lambda$$

e, como  $\langle \tilde{U}(i, \cdot), u \rangle = 0$ ,

$$\tilde{U}^H A u = \tilde{U}^H \lambda u = \lambda \tilde{U}^H u = \lambda 0 = 0, \text{ assim}$$

$$U^H A U = \begin{pmatrix} \lambda & u^H A \tilde{U} \\ 0 & \tilde{U}^H A \tilde{U} \end{pmatrix}, \text{ e definimos}$$

$$U^H A U = \begin{pmatrix} \lambda & \tilde{A}_{12} \\ 0 & \tilde{A}_{22} \end{pmatrix}$$

Pela hipótese de indução, existe uma matriz unitária  $P$ , tal que  $P^H \tilde{A}_{22} P = \tilde{T}$  e  $\tilde{T}$  é uma matriz triangular superior. Então,

$$\tilde{A}_{22} = P \tilde{T} P^H, \text{ logo}$$

$$U^H A U = \begin{pmatrix} \lambda & \tilde{A}_{12} \\ 0 & P \tilde{T} P^H \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} \begin{pmatrix} \lambda & \tilde{A}_{12} P \\ 0 & \tilde{T} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & P^H \end{pmatrix}. \text{ Assim}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & P^H \end{pmatrix} U^H A U \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} = \begin{pmatrix} \lambda & \tilde{A}_{12} P \\ 0 & \tilde{T} \end{pmatrix} = T$$

Portanto, a matriz  $T$  é triangular superior e  $Q = U \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix}$  é matriz unitária.

## 6.2 O Teorema Espectral

**Teorema 6.2.1 (Teorema Espectral).** *Seja  $A$  uma matriz normal, então existe uma matriz unitária  $U$  que diagonaliza  $A$ .*

**Prova:**

Provamos no teorema 6.1.1 a existência de uma matriz unitária  $U$  tal que  $U^H A U = T$ , em que  $T$  é triangular superior.

Então:

$$T^H = (U^H A U)^H = U^H A^H (U^H)^H = U^H A^H U$$

Como  $A$  é normal ( $A^H A = A A^H$ ),  $T^H T = T T^H$ .

Logo,  $T$  é diagonal.

**Observação:** No livro de LEON (5), a prova deste teorema é vista com maior detalhamento.

**Corolário 6.2.1.** *Se  $A$  é uma matriz real e simétrica, então existe uma matriz ortogonal  $U$  que diagonaliza  $A$ , isto é,  $U^T A U = D$ , em que  $D$  é diagonal.*

**Prova:**

**Resultados auxiliares:** Seja  $A \in \mathbb{R}^{n \times n}$  simétrica. Temos:

$$(i) A = V D V^H \Rightarrow A^T = \bar{V} D V^T = \overline{V \bar{D} \bar{V}^T} = \overline{V \bar{D} V^H}$$

$$A = A^T \Rightarrow V D V^H = \overline{V \bar{D} V^H}$$

$$\text{Como } A \in \mathbb{R}^{n \times n}, V \bar{D} V^H = \overline{V \bar{D} V^H} = V D V^H \Rightarrow D = \bar{D}$$

Logo, os autovalores de  $A$ , simétrica real, são reais.

(ii)  $V = B + iC$  é inversível então  $B, C$  são inversíveis.

**Demonstração**



$V$  é inversível, então a única solução para  $Vx = 0$  é  $x = 0$ . Mas, se  $x \in \mathbb{R}^n$ ,  $Vx = 0 \Leftrightarrow Bx = Cx = 0$ .

Logo, a única solução real, tanto para  $Bx = 0$  como para  $Cx = 0$ , é a solução  $x = 0$ , ou equivalentemente,  $B$  e  $C$  são inversíveis.

$$(iii) \quad \begin{cases} A \in \mathbb{R}^{n \times n} \\ AV = VD, \quad D \in \mathbb{R}^{n \times n} \\ V = B + iC \end{cases} \Rightarrow \begin{cases} AB = BD \\ AC = CD \end{cases}$$

(iv) Se  $Av = \lambda v$ , em que  $v \in \mathbb{R}^n$  e  $v \neq 0$ ; e,  $Aw = \mu w$ , em que  $w \in \mathbb{R}^n$ ,  $w \neq 0$  e  $\lambda \neq \mu$ , então  $v^T w = 0$ .

### Demonstração

$$v^T Aw = v^T \mu w = \mu v^T w$$

$$(v^T Aw)^T = w^T A^T v = w^T Av = w^T \lambda v = \lambda w^T v = \lambda v^T w$$

Como  $v^T Aw = (v^T Aw)^T$  e  $\lambda \neq \mu$ , então  $v^T w = 0$

**Prova do corolário:** Pelo teorema espectral existe uma matriz unitária  $V$  e uma matriz diagonal  $D$  tais que  $A = VDV^H$ . Por (i),  $D \in \mathbb{R}^{n \times n}$ . Por (ii) e (iii), existe  $B \in \mathbb{R}^{n \times n}$  inversível tal que  $AB = BD$ . Vamos supor, sem perda de generalidade, que diagonal de  $D$  está ordenada de tal modo que  $d_{11} \geq d_{22} \geq \dots \geq d_{nn}$ , correspondente aos autovalores  $\lambda_1 > \lambda_2 > \dots > \lambda_q$ .

Sejam  $B_1 = B(:, 1 : r_1)$ ,  $B_2 = B(:, r_1 + 1 : r_2)$ , ...,  $B_q = B(:, r_{q-1} + 1 : r_q)$  as matrizes colunas de  $B$  que são autovetores respectivamente de  $\lambda_1, \lambda_2, \dots, \lambda_q$ .

Sejam  $B_1 = Q_1 R_1, B_2 = Q_2 R_2, \dots, B_q = Q_q R_q$  as decomposições QR de  $B_1, B_2, \dots, B_q$ , respectivamente ( $Q_i$  é da mesma ordem que  $B_1$ ). É claro que  $Q_1, Q_2, \dots, Q_q$  são matrizes formadas por autovalores de  $A$  associados respectivamente a  $\lambda_1, \lambda_2, \dots, \lambda_q$ . Sabemos, também, por (iii), que  $Q_j^T Q_i = 0$  para  $i \neq j$  (autovetores associados a autovalores distintos são ortogonais). Seja  $Q = \begin{pmatrix} Q_1 & \dots & Q_q \end{pmatrix}$ ,  $Q$  é ortogonal e

$$AQ = \begin{pmatrix} AQ_1 & \dots & AQ_q \end{pmatrix} = \begin{pmatrix} \lambda_1 Q_1 & \dots & \lambda_q Q_q \end{pmatrix} = \begin{pmatrix} Q_1 & \dots & Q_q \end{pmatrix} \begin{pmatrix} d_{11} & & \\ & \ddots & \\ & & d_{qq} \end{pmatrix} = QD$$

## 7 Decomposição em Valores Singulares

### 7.1 Os Valores Singulares de uma Matriz

Seja  $A \in \mathbb{R}^{m \times n}$ . Então  $A^T A$  é uma matriz simétrica, que pode ser diagonalizada por uma matriz ortogonal. Seja  $\alpha = \{v_1, v_2, v_3, \dots, v_n\}$  uma base ortonormal de  $\mathbb{R}^n$  formada pelos autovetores de  $A^T A$ , e sejam  $\lambda = \{\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n\}$  os respectivos autovalores. Então

$$\|Av_i\|^2 = (Av_i)^T Av_i = v_i^T A^T Av_i$$

Como  $v_i$  é autovetor de  $A^T A$

$$\|Av_i\|^2 = v_i^T \lambda_i v_i$$

Uma vez que  $v_i$  é unitário

$$\|Av_i\|^2 = \lambda_i$$

Assim sendo, os autovalores de  $A^T A$  são todos não negativos.

Os valores singulares de  $A$  são as raízes quadradas dos autovalores de  $A^T A$  reordenados de forma decrescente e, denotados por  $\sigma_i$ . Como visto acima, são os comprimentos dos vetores  $Av_i$ . O exemplo a seguir ilustra de forma geométrica um valor singular.

- **Exemplo** Seja  $A = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}$  uma matriz e  $\|u\| = 1$  um vetor unitário, então a transformação linear  $T(u) = A(u)$  leva a esfera unitária de  $\mathbb{R}^3$  sobre uma elipse em  $\mathbb{R}^2$  (teorema 7.2.2). Encontre um vetor unitário  $u$  no qual o comprimento  $\|Au\|$  seja máximo e calcule esse comprimento.

**Resolução:**  $\|Au\|$  atinge seu máximo para o mesmo  $u$  que torna máximo  $\|Au\|^2$ . Então,

$$\|Au\|^2 = (Au)^T (Au) = u^T A^T Au = u^T (A^T A)u$$

Como  $A^T A$  é uma matriz simétrica, o problema passa a ser a maximização da forma

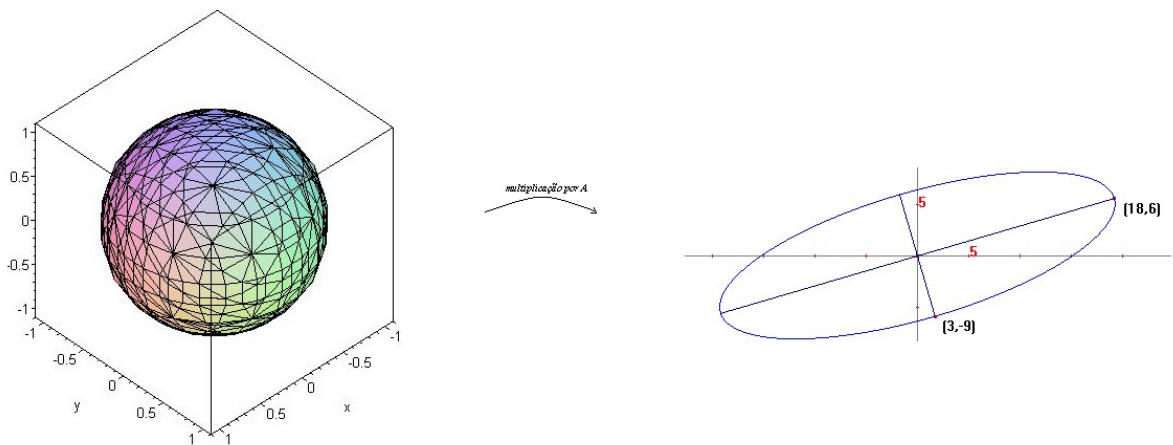


Figura 5: Visualização de transformação de  $\mathbb{R}^3$  em  $\mathbb{R}^2$ .

quadrática  $u^T(A^T A)u$  sujeita a restrição  $\|u\|_2 = 1$ . O valor máximo será então um autovalor  $\lambda_i$  de  $A^T A$  e será atingido em um autovetor unitário de  $A^T A$  associado a  $\lambda_i$  (teorema 7.2.2).

$$A^T A = \begin{pmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{pmatrix} \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix} = \begin{pmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{pmatrix}$$

Os autovalores de  $A^T A$  são  $\lambda_1 = 360$ ,  $\lambda_2 = 90$  e  $\lambda_3 = 0$ . Autovetores associados são respectivamente

$$v_1 = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix}, \quad v_2 = \begin{pmatrix} -\frac{2}{3} \\ -\frac{1}{3} \\ \frac{2}{3} \end{pmatrix}, \quad v_3 = \begin{pmatrix} \frac{2}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{pmatrix}$$

O valor máximo de  $\|Au\|^2$  é 360, atingido quando  $u$  é o vetor unitário  $v_1$ . O vetor  $Av_1$  é o ponto na elipse, ilustrada na figura 7.1, mais afastado possível da origem, isto é,

$$Av_1 = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix} \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix} = \begin{pmatrix} 18 \\ 6 \end{pmatrix}$$

Para  $\|u\| = 1$ , o valor máximo de  $\|Au\|$  é  $\|Av_1\| = \sqrt{360} = 6\sqrt{10}$ .

Como os autovalores de  $A^T A$  são 360, 90 e 0, os valores singulares de  $A$  são

$$\sigma_1 = 6\sqrt{10}, \quad \sigma_2 = 3\sqrt{10}, \quad \text{e} \quad \sigma_3 = 0.$$

O primeiro valor singular de  $A$  é o máximo de  $Au$  sobre todos os vetores unitários, e o máximo é atingido no autovetor unitário  $v_1$ ; o segundo valor singular de  $A$  é o máximo de  $Au$  sobre todos os vetores ortogonais a  $v_1$ , e esse máximo é atingido no segundo autovetor unitário  $v_2$

$$Av_2 = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix} \begin{pmatrix} -\frac{2}{3} \\ -\frac{1}{3} \\ \frac{2}{3} \end{pmatrix} = \begin{pmatrix} 3 \\ -9 \end{pmatrix}$$

Este ponto está no eixo menor da elipse da figura 7.1. Notemos que os dois primeiros valores singulares de  $A$  são os comprimentos dos semi-eixos maior e menor da elipse.

Uma das utilidades computacionais dos valores singulares é a análise do posto de uma matriz.

*Em alguns casos (...). O método óbvio de contar o número de colunas pivô em  $A$  não funciona bem se  $A$  for reduzida por linhas por um computador. (...) Na prática, a maneira mais confiável de estimar o posto de uma matriz  $A$  grande é contar o número de valores singulares não-nulos. Nesse caso, valores singulares extremamente pequenos são considerados nulos para todos os efeitos, e o posto efetivo da matriz é o número obtido contando-se os valores singulares não-nulos restantes. David C. LAY(4).*

**Teorema 7.1.1.** *Seja  $\alpha = \{v_1, v_2, v_3, \dots, v_n\}$  uma base ortonormal para  $\mathbb{R}^n$  consistindo de autovetores de  $A^T A$  ordenados de forma que os autovalores correspondentes estejam em ordem decrescente. Suponhamos que  $A$  tenha  $p$  valores singulares não nulos. Então  $\beta = \{Av_1, Av_2, Av_3, \dots, Av_p\}$  é uma base ortogonal para o espaço coluna de  $A$  e posto de  $A = p$ .*

**Demonstração:** Como  $v_i$  e  $v_j$  são ortogonais para qualquer  $i \neq j$ , temos

$$\langle Av_i, Av_j \rangle = (Av_i)^T Av_j = v_i^T A^T Av_j = v_i^T (\lambda_j v_j) = 0$$

Logo  $Av_1, Av_2, Av_3, \dots, Av_p$  é um conjunto ortogonal. Como os valores singulares de  $A$  são os comprimentos dos vetores  $Av_i$  e existem somente  $p$  valores singulares não-nulos,

temos que  $Av_i \neq 0$  se e somente se  $1 \leq i \leq p$ . Portanto, o conjunto  $\beta$  é composto por vetores  $LI$  pertencentes ao espaço coluna de  $A$ , e, qualquer que seja  $w$  pertencente ao espaço coluna de  $A$ ,

$$\begin{aligned}w = Au &= c_1Av_1 + c_2Av_2 + \dots + c_pAv_p + \dots + c_nAv_n \\ &= c_1Av_1 + c_2Av_2 + \dots + c_pAv_p + 0 \dots + 0\end{aligned}$$

Concluimos que  $w \in [Av_1, Av_2, \dots, Av_p]$ , mostrando que  $\beta$  é uma base para o espaço coluna de  $A$ . Portanto, posto de  $A$  é igual a  $p$ .

## 7.2 A Decomposição em Valores Singulares - o Teorema

A decomposição de uma matriz  $A$  em valores singulares envolve uma matriz  $\Sigma_{m \times n}$  da forma

$$\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \quad (7.1)$$

em que  $D = \text{diag}(\sigma_i)$  é uma matriz diagonal  $p \times p$  para algum  $p$  não excedendo o menor valor entre  $m$  e  $n$ , ou seja, temos  $m - p$  linhas de “zeros” e  $n - p$  colunas de “zeros” em  $\Sigma$ .

Qualquer fatoração  $A = U\Sigma V^T$ , com  $U$  e  $V$  ortogonais,  $\Sigma$  como em 7.1 com elementos diagonais positivos em  $D$ , é chamada uma decomposição em valores singulares (DVS) de  $A$ . No caso,  $U$  é  $m \times m$  e  $V$  é  $n \times n$ , e  $\Sigma$  é retangular, com a mesma dimensão de  $A$ . As entradas da diagonal de  $\Sigma$ , que são  $\Sigma_{ii} = \sigma_i$ , são ordenadas de forma decrescente. Em uma tal decomposição, as colunas de  $U$  são chamadas de vetores singulares à esquerda de  $A$ , e as colunas de  $V$  são chamadas de vetores singulares à direita de  $A$ .

A DVS está intimamente relacionada à diagonalização de matrizes simétricas. Dada uma matriz  $A \in \mathbb{R}^{n \times n}$  simétrica, existe uma matriz ortogonal  $V$  e uma matriz diagonal  $D$  tal que  $A = VDVT$ . As colunas de  $V$  são os autovetores de  $A$  e formam uma base ortonormal para  $\mathbb{R}^n$ ; as entradas da diagonal de  $D$  são os autovalores de  $A$ .

Agora, analisemos a DVS de uma matriz  $A_{m \times n}$ . A matriz  $\Sigma$  corresponde à matriz da transformação de  $\mathbb{R}^n$  em  $\mathbb{R}^m$ , em relação às bases dadas pelas colunas de  $V$  e  $U$ . Quando elas são usadas para representar vetores no domínio, a natureza da transformação se torna evidente: ela é simplesmente a dilatação de alguns componentes e a contração de outros, de acordo com a magnitude dos valores singulares, conforme já vimos no exemplo 7.1. Há possibilidade de descartar componentes ou acrescentar zeros, uma necessidade conforme a mudança de dimensão. Vista desta perspectiva, a DVS nos diz como escolher uma base ortonormal, de forma que, a transformação é representada por uma matriz simples, isto é, uma matriz diagonal.

**Teorema 7.2.1 (Teorema da DVS).** *Seja  $A_{m \times n}$  uma matriz de posto  $p$ . Então existe uma matriz  $\Sigma_{m \times n}$ , em que os elementos diagonais de  $D$  são os  $p$  primeiros valores singulares de  $A$ ,  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p > 0$ , e existem uma matriz ortogonal  $U_{m \times m}$  e uma matriz ortogonal  $V_{n \times n}$  tais que*

$$A = U\Sigma V^T$$

**Demonstração:** Seja  $\lambda_i$  e  $v_i$  como no Teorema 7.1.1, de modo que  $\{Av_1, \dots, Av_p\}$  seja uma base ortogonal para o espaço coluna de  $A$ . Normalizamos cada  $Av_i$  para obter uma base ortonormal  $\alpha = \{u_1, \dots, u_p\}$ , em que

$$u_i = \frac{1}{\|Av_i\|} Av_i = \frac{1}{\sigma_i} Av_i$$

e

$$Av_i = \sigma_i u_i \quad (1 \leq i \leq p)$$

Vamos, agora, estender  $\alpha$  a uma base ortonormal  $\{u_1, \dots, u_m\} \in \mathbb{R}^m$  e definir as matrizes

$$U = \begin{pmatrix} u_1 & u_2 & \dots & u_m \end{pmatrix} \quad \text{e} \quad V = \begin{pmatrix} v_1 & v_2 & \dots & v_n \end{pmatrix}$$

Por construção,  $U$  e  $V$  são matrizes ortogonais. Além disso

$$AV = \begin{pmatrix} Av_1 & \dots & Av_p & 0 & \dots & 0 \end{pmatrix} = \begin{pmatrix} \sigma_1 u_1 & \dots & \sigma_p u_p & 0 & \dots & 0 \end{pmatrix}$$

Seja  $D$  a matriz, de existência garantida pelo teorema 6.2.1, cujos elementos diagonais são  $\sigma_1, \dots, \sigma_p$  e seja  $\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$ , como definida na equação 7.1. Então

$$U\Sigma = \begin{pmatrix} u_1 & u_2 & \dots & u_m \end{pmatrix} \begin{pmatrix} \sigma_1 & & & & & 0 \\ 0 & \sigma_2 & & & & 0 \\ & & \ddots & & & \\ & & & & \sigma_p & \\ 0 & 0 & & & & 0 \end{pmatrix}$$

isto é,

$$\begin{aligned} U\Sigma &= \begin{pmatrix} \sigma_1 u_1 & \dots & \sigma_p u_p & 0 & \dots & 0 \end{pmatrix} \\ &= AV \end{aligned}$$

Como  $V$  é uma matriz ortogonal,  $U\Sigma V^T = AVV^T = A$ .

**Corolário 7.2.1.** *Seja  $A$  uma matriz  $m \times n$ , em que  $n \leq m$ . Escrevendo-se  $U = (u_1 \ u_2 \ \dots \ u_n)$  e  $V = (v_1 \ v_2 \ \dots \ v_n)$ , temos que*

$$A = U\Sigma V^T = \sum_{i=1}^n \sigma_i u_i v_i^T \quad (\text{somatório de matrizes de posto 1.})$$

*Então a matriz de posto  $p \leq n$  mais perto de  $A$  é a matriz  $A_p = \sum_{i=1}^p \sigma_i u_i v_i^T$ , cuja distância até  $A$  é  $\|A - A_p\| = \sigma_{p+1}$ .*

**Teorema 7.2.2.** *Sejam  $A = U\Sigma V^T$ , a DVS de uma matriz  $A_{m \times n}$ , e  $S$  uma esfera unitária em  $\mathbb{R}^n$ :  $S = \{u \in \mathbb{R}^n \text{ tal que } \|u\|_2 = 1\}$ . Seja, também,  $AS$  a imagem de  $S$  por  $A$ :  $AS = \{Au \text{ tal que } u \in \mathbb{R}^n \text{ e } \|u\|_2 = 1\}$ . Então  $AS$  é um elipsóide centrado na origem de  $\mathbb{R}^m$ , com eixos  $\sigma_i v_i$ , se posto de  $A$  igual a  $n$ .*

**Demonstração:** Para a prova, nós construiremos a transformação  $AS$  trabalhando com a multiplicação fator por fator, de  $A = U\Sigma V^T$ .

Visto que  $V$  é ortogonal,  $V$  transforma um vetor unitário em outro vetor unitário, então  $V^T S = S$ .

Como  $u \in S$  se e somente se  $\|u\|_2 = 1$ ,  $v \in \Sigma S$  se e somente se  $\sum_{i=1}^n \left(\frac{v_i}{\sigma_i}\right)^2 = 1$ . Vemos que esta última igualdade define um elipsóide com eixos  $\sigma_i e_i$  em que  $e_i$  é  $i$ -ésimo vetor canônico (no espaço  $x_{n+1} = \dots = x_m = 0$ ). Finalmente, multiplicando-se cada  $v = \Sigma u$  por  $U$  rotacionamos o elipsóide de forma que cada  $e_i$  torna-se  $u_i$ , a  $i$ -ésima coluna de  $U$ . Conseqüentemente o comprimento do maior eixo do elipsóide será o maior  $\sigma_i$ .

Quando  $\Sigma$  contém linhas ou colunas nulas, é possível obter uma decomposição mais compacta de  $A$ . Seja  $p$  o posto da matriz  $A$ . Particionamos  $U$  e  $V$  em submatrizes de modo que o primeiro bloco de cada uma contenha  $p$  colunas:

$$U = \begin{pmatrix} U_p & U_{m-p} \end{pmatrix} \quad \text{em que} \quad U_p = \begin{pmatrix} u_1 & \dots & u_p \end{pmatrix}$$

$$V = \begin{pmatrix} V_p & V_{n-p} \end{pmatrix} \quad \text{em que} \quad V_p = \begin{pmatrix} v_1 & \dots & v_p \end{pmatrix}$$

Nesse caso,  $U$  é  $m \times p$  e  $V$  é  $n \times p$ . A multiplicação de matrizes em bloco nos mostra,



então, que

$$A = \begin{pmatrix} U_p & U_{m-p} \end{pmatrix} \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_p^T \\ V_{n-p}^T \end{pmatrix} = U_p D V_p^T$$

Essa fatoração é conhecida como decomposição em valores singulares reduzida.

Resumindo, uma matriz  $A \in \mathbb{R}^{m \times n}$  pode ser expressa por um produto  $U\Sigma V^T$ , quando  $U$  e  $V$  são matrizes ortogonais e  $\Sigma$  é uma matriz diagonal. A matriz  $V$  é obtida da fatoração diagonal de  $A^T A = V D V^T$ ,  $U$  provém da normalização da imagem de  $A$  sobre  $V$  estendida a uma base ortonormal, se necessário, conforme já descrito. As matrizes  $U$  e  $V$  não estão unicamente determinadas por  $A$ , mas os elementos diagonais de  $\Sigma$  são necessariamente os valores singulares de  $A$ .

**Exemplo** Encontrar uma decomposição em valores singulares para  $A = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}$ .

**Resolução:** No exemplo 7.1 já encontramos  $v_1, v_2$ , e  $v_3$  que são os nossos vetores singulares à direita de  $A$ . Também, tínhamos calculado  $Av_1$  e  $Av_2$ . Agora, definimos

$$u_1 = \frac{1}{\sigma_1} Av_1 = \frac{1}{6\sqrt{10}} \begin{pmatrix} 18 \\ 6 \end{pmatrix} = \begin{pmatrix} \frac{3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{pmatrix}$$

$$u_2 = \frac{1}{\sigma_2} Av_2 = \frac{1}{3\sqrt{10}} \begin{pmatrix} 3 \\ -9 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{10}} \\ \frac{-3}{\sqrt{10}} \end{pmatrix}$$

Logo  $\{u_1, u_2\}$  é uma base para  $\mathbb{R}^2$ . Sejam  $U = \begin{pmatrix} u_1 & u_2 \end{pmatrix}$  e  $V = \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix}$  e definindo

$$D = \begin{pmatrix} 6\sqrt{10} & 0 \\ 0 & 3\sqrt{10} \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{pmatrix}$$

Portanto, a DVS de  $A$  é

$$A = \begin{pmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{-3}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

## 8 *Compressão de Imagem Usando a DVS*

A DVS tem uma aplicação natural: a **compressão de imagens digitais**. Uma imagem digital é uma matriz  $m \times n$ , em que as entradas são traduções numéricas da luminosidade e da cor dos “patches” (a tela fica dividida em  $m \times n$  patches, como uma colcha de retalhos, que é a tradução literal de patches). No caso de imagem colorida podemos ter uma matriz  $m \times n \times 3$ , em que o primeiro elemento é o tom de vermelho, o segundo o tom de verde e, o terceiro o tom de azul.

Se  $A = U\Sigma V^T$  é a DVS de  $A$ ,  $A = \sum_{i=1}^p \sigma_i u_i v_i^T$ , em que  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p > 0$  são os valores singulares de  $A$  e, para todo  $i$ ,  $u_i$   $v_i$  são as  $i$ -ésimas colunas de  $U$  e  $V$ , respectivamente, queremos encontrar uma aproximação para  $A$  que capture a maior parte dos significados dos dados, no caso, acuidade visual e relativa perfeição da imagem. A aproximação por DVS diminui ou elimina certa redundância de dados. O posto de uma matriz especifica o número de linhas ou colunas linearmente independentes da matriz, ou seja, ele é um mensurador da redundância. Uma matriz que tenha um posto  $p$  pequeno é altamente redundante, pois necessitam-se apenas  $p$  colunas de  $U$  e  $V$ , mais  $p$  valores singulares para descrever exatamente  $A$ . Ou seja,  $pm + np = p(m + n)$  elementos.

Então com a DVS podemos expressar os dados muito mais eficientemente do que simplesmente por uma listagem de todas as entradas da matriz. Como um exemplo, suponhamos que um scanner é usado para digitalizar uma fotografia substituindo a imagem por uma matriz de  $m \times n$  pixels, na qual cada nível de cinza é representado numa escala de 0 a 1. Em qualquer escala, as características de uma imagem são traduzidas redundantemente nas linhas ou colunas da matriz. Nós podemos rever estas características de forma menos redundante, através da aproximação da matriz original por uma matriz de posto menor do que o mínimo entre  $m$  e  $n$ .

Temos do corolário 7.2.1 que  $A_p = \sum_{i=1}^p \sigma_i u_i v_i^T$  é a melhor aproximação de  $A$  com posto  $p$ , no sentido de minimizar  $\|A - A_p\| = \sigma_{p+1}$ . Lembrando que somente com  $mp + np = (m + n)p$  elementos, podemos estocar  $\{u_1, \dots, u_p\}$  e  $\{\sigma_1 v_1, \dots, \sigma_p v_p\}$ , pois estamos guardando as matrizes  $U_{m \times p}$  e  $(\Sigma V^T)_{p \times n}$ . Normalmente, utilizaríamos  $mn$  elementos para estocar  $A$ , que é um número muito grande comparado a  $(m + n)p$ , se  $p$  é pequeno em relação a  $m$  ou a  $n$ . Conseqüentemente, o erro relativo da aproximação de  $A$  por  $A_p$  é medido por  $\frac{\sigma_{p+1}}{\sigma_p}$  e a razão de compressão é medida por  $\frac{(m + n)p}{mn}$ .

Consideremos alguns aspectos de suma importância ao entendimento dos exemplos:

- Na obtenção da DVS utilizamos o MatLab<sup>1</sup>, que também foi utilizado na criação e compressão das imagens, vista a sua polivalência.
- O MatLab apresenta diversos mapas de cores predefinidos, entre eles as opções mapa default (figura 6) e mapa gray (figura 7), e outros podem ser criados. A matriz que armazena um mapa de cores é uma matriz  $C_{m \times 3}$ , em que  $C$  terá um número qualquer de linhas (até 64) e necessariamente três colunas. Cada linha da matriz é interpretada como uma cor, em que o primeiro elemento determina a intensidade de vermelho, o segundo de verde e o terceiro de azul. A intensidade das cores pode ser especificada no intervalo  $[0, 1]$ . Por exemplo,  $\begin{pmatrix} 0 & 0 & 0 \end{pmatrix}$  representa preto,  $\begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$  especifica branco,  $\begin{pmatrix} 1 & 0 & 0 \end{pmatrix}$  é vermelho,  $\begin{pmatrix} 0.5 & 0.5 & 0.5 \end{pmatrix}$  é cinza, e  $\begin{pmatrix} 127/255 & 1 & 212/255 \end{pmatrix}$  é aquamarinha.
- As imagens neste trabalho foram geradas com os comandos `image`, `imagesc` e o comando `pcolor`.

O comando `image` interpreta, somente, valores inteiros entre 1 e 64 procedendo arredondamento (o menor inteiro mais próximo de), para números racionais entre 1 e 64; ou tratando como 1 e como 64, números abaixo de 1 ou acima de 64, respectivamente. Quando a matriz  $A$  que representa a imagem é uma matriz  $m \times n$ , os elementos de  $A$  são usados como índices do corrente mapa de cores para determinar as cores, ou seja, 1 corresponde a azul, 32 a aquamarinha, 64 a vermelho e assim por diante, no mapa `default` (ver figura 6). Porém, quando  $A$  é uma matriz  $m \times n \times 3$  os elementos em  $A(:, :, 1)$  são interpretados como as intensidades de vermelho, em  $A(:, :, 2)$  como as intensidades de verde e em  $A(:, :, 3)$  como as de

---

<sup>1</sup>The MathWorks, Inc. All Rights Reserved.

azul, independentemente do mapa de cores utilizado. Na figuras abaixo podemos ver o comando `image(A)` aplicado a uma matriz  $A_{1 \times 64}$  com entradas contendo inteiros do 1 ao 64.

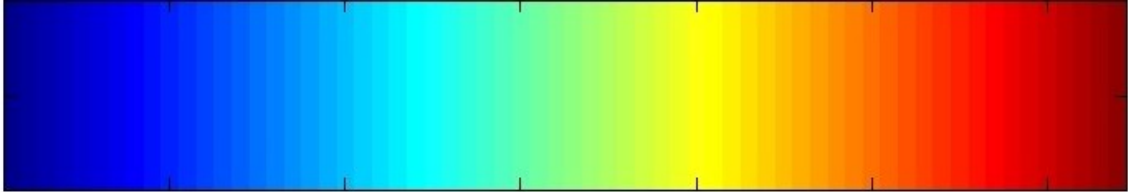


Figura 6: Cores de  $A$  no mapa `default`.

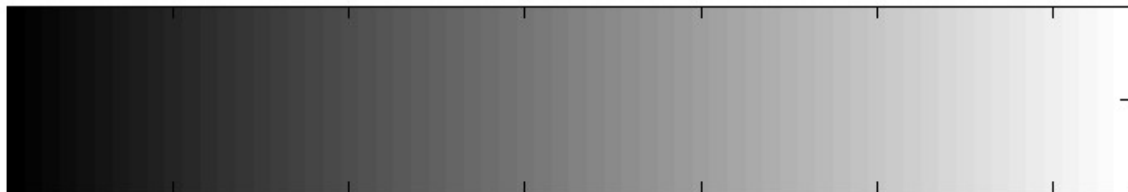


Figura 7: Cores de  $A$  no mapa `gray`.

Já o comando `imagesc`, escalona as cores do corrente mapa de cores conforme as entradas da matriz  $A$ . Por exemplo, se  $A = \begin{pmatrix} 0.2 & 0.64 & 1.28 \end{pmatrix}$  as cores, no corrente mapa de cores, serão interpretadas com 0.2 equivalendo a 1; 0.64 equivalendo a 32; e, 1.28 equivalendo a 64; do comando `image`. Quando  $A$  é  $m \times n \times 3$  a imagem é gerada da mesma forma que no comando `image`.

- Com a utilização do comando `pcolor` podemos notar, claramente, cada célula que compõe a imagem, pois, as células ficam delimitadas por sombras (ver figura 8).
- Quando importamos uma imagem ao MatLab ele a armazena em uma matriz (nome do arquivo) $_{m \times n \times 3}$  no formato `uint8`; porém, para realizarmos cálculos devemos converter a matriz para o formato `double`; e, quando desejarmos visualizar a imagem devemos reconverter a matriz ao formato `uint8`; além disso, o formato `uint8` ocupa menos bytes do que `double`.
- Por fins computacionais, o MatLab considera como sendo zero todos os valores singulares muito pequenos, e os anota com quatro casas decimais 0.0000; e, para aqueles valores singulares que são realmente zero, anota simplesmente 0.

### Exemplos de imagens e suas respectivas compressões:

1. Considere a imagem da figura 8. Esta imagem tem  $20 \times 20$  pixels, o que corresponde a uma matriz  $M_{20 \times 20}$ , em que cada entrada é exatamente a 0 ou 1. Na imagem criada, cada célula está colorida em preto (correspondendo a entrada 0 da matriz) ou branco (quando a entrada é 1). O mapa de cores utilizado contém apenas as cores preto e branco, ou seja, é definido pela matriz  $mapa = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$ . Abaixo estão descritos todos os 20 valores singulares de  $M$  com aproximação de 4 casa decimais.

9.2350	3.7440	3.3476	2.5177	2.1369	1.7886	1.4066
1.0273	0.5938	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0	0	0	

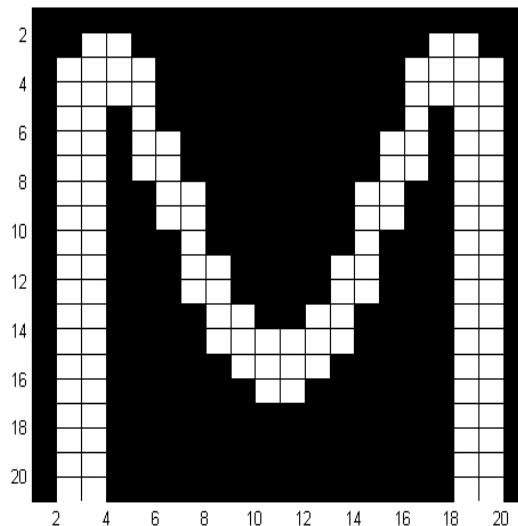


Figura 8: Uma imagem  $20 \times 20$

Seja  $U\Sigma V^T = M$  a decomposição em valores singulares de  $M$ . As próximas figuras mostram aproximações (compressões) feitas da figura 8 pelo método da DVS. Para tal, foi usada no MatLab a seguinte rotina:

```
[u s v]=svd(M);
M1=(u(:,1:p)*s(1:p,1:p)*v(:,1:p)')
pcolor(M1), view([0 270])
```

Fazendo-se alterações no posto da matriz de aproximação  $M1$  mudamos a razão de compressão e conseqüente a qualidade da imagem. Lembrando-se que  $M$  tem 9 valores singulares não nulos, ou seja, posto de  $M$  é igual a 9.

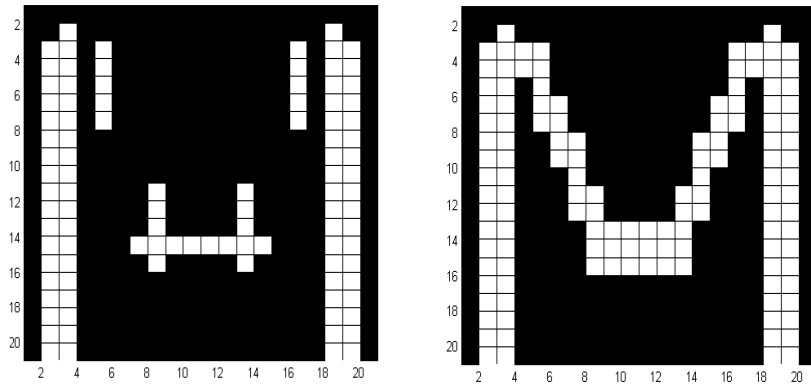


Figura 9: Compressão da figura 8 com  $p = 2$  (esquerda)  $p = 3$  (direita)

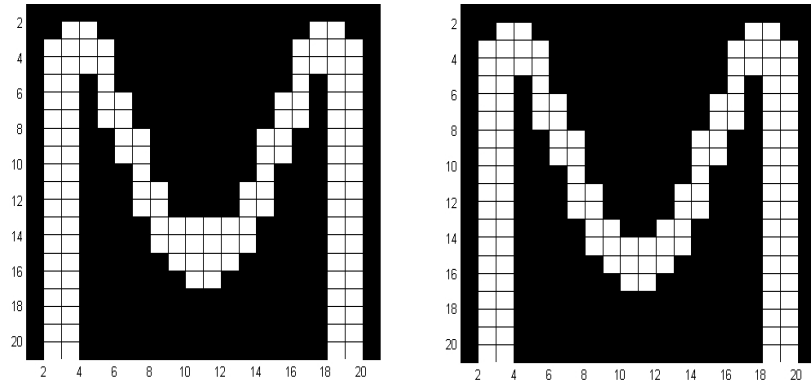


Figura 10: Compressão da figura 8 com  $p = 4$  (esquerda)  $p = 5$  (direita)

Com  $p = 4$  praticamente reconstruímos a imagem e com  $p = 5$  não se pode notar nenhuma diferença de acuidade entre a imagem original (figura 8) e a imagem acima (direita).

Medindo a qualidade da compressão:

$p$	Erro relativo $\frac{\sigma_{p+1}}{\sigma_p}$	Razão de compressão $\frac{1}{10}p$
2	0.8941	$\frac{1}{5} = 0.2$
3	0.7521	$\frac{3}{10} = 0.3$
4	0.8487	$\frac{2}{5} = 0.4$
5	0.8370	$\frac{1}{2} = 0.5$
9	0.0000	$\frac{9}{10} = 0.9$

2. Para este exemplo, trabalharemos com uma imagem um pouco mais complexa, que está associada a uma matriz  $B_{488 \times 640}$  cujo posto é  $p = 54$ .

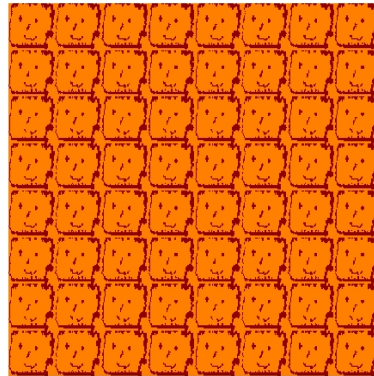


Figura 11: Uma imagem  $576 \times 1280$

Faremos a compressão desta imagem e visualizaremos os resultados com as seguintes linhas de comando:

```
[u s v]=svd(B);
colormap('default')
B1=(u(:,1:p)*s(1:p,1:p)*v(:,1:p)')
image(B1)
```

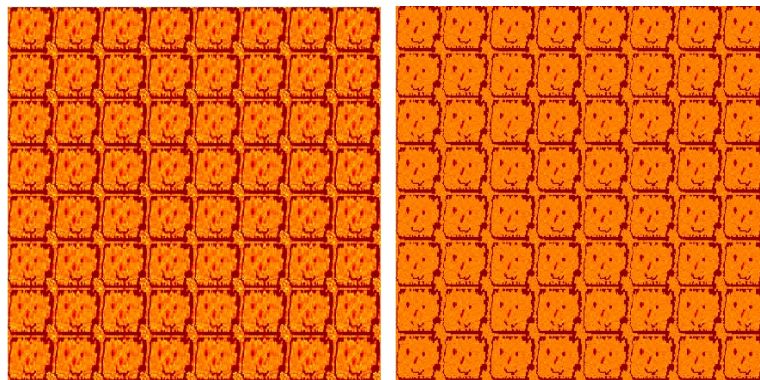


Figura 12: Compressão da figura 11 com  $p = 10$  (esquerda) e com  $p = 20$  (direita)

A qualidade da imagem não sofre melhorias significativas a partir de  $p = 20$  até  $p = 40$ .

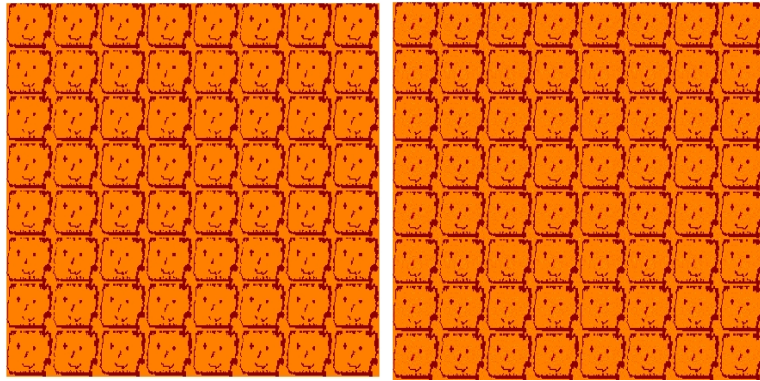


Figura 13: Imagem original (esquerda) e sua compressão com  $p = 40$  (direita)

Quando  $p = 40$ , podemos verificar que a imagem produzida é, praticamente, a própria imagem original, conforme pode ser visto na figura acima. A razão de compressão é  $\frac{141}{39040}p$  e, portanto para  $p = 40$  temos uma razão de compressão igual a  $\frac{141}{976} \cong 0,15$ , ou seja, conseguimos uma compressão na ordem de 85%.



3. Para um ultimo exemplo, trabalharemos com a foto do Departamento de Matemática da UFSC. Tal foto é importada do formato `jpg` para o ambiente MatLab onde fica armazenada como uma matriz  $A_{242 \times 425 \times 3}$ , ou seja, como três matrizes  $A_{242 \times 425}$ , em que a primeira matriz armazena as intensidades de vermelho, a segunda de verde e a terceira de azul. As três matrizes que formam a foto  $A1_{242 \times 425}$ ,  $A2_{242 \times 425}$  e  $A3_{242 \times 425}$  tem o mesmo posto  $p = 242$ . Podemos ver a imagem original em `jpg` (esquerda) e a sua representação feita pelo no MatLab (direita).



Figura 14: Foto do Departamento em `jpg` (esquerda) e no formato `mat` (direita)

A foto originalmente em `jpg` tem tamanho de 71.031 bytes, no formato `bmp` a mesma coresponde a 311.296 bytes, enquanto que, no formato `mat`, do MatLab, a matriz  $A$  associada a essa foto tem 308.550 bytes. Procederemos a aproximação de  $A$  por matrizes de posto menor, conseqüentemente diminuiremos o número de bytes necessários para armazená-la. Para isso, trabalharemos com a seguinte rotina:

```
a=double(A); x1=a(:,:,1); x2=a(:,:,2); x3=a(:,:,3);
[u1,s1,v1]=svd(x1); [u2,s2,v2]=svd(x2); [u3,s3,v3]=svd(x3);
t(:,:,1)=u1(:,1:p)*s1(1:p,1:p)*v1(:,1:p)';
t(:,:,2)=u2(:,1:p)*s2(1:p,1:p)*v2(:,1:p)';
t(:,:,3)=u3(:,1:p)*s3(1:p,1:p)*v3(:,1:p)'; mtm=(uint8(t));
image(mtm), axis off, axis image
```

Abaixo, podemos ver as compressões das fotos feitas com  $p = 10$  e  $p = 50$ , respectivamente.



Figura 15: Compressão da figura 14 com  $p = 10$  (esquerda) e com  $p = 50$  (direita)

Observe que com  $p = 50$  a imagem obtida é muito semelhante à original. Podemos ver na figura 16 abaixo, a comparação entre a imagem original e uma imagem comprimida com  $p = 100$ .



Figura 16: Imagem original (esquerda) e sua compressão com  $p = 100$  (direita)

Agora, armazenamos as matrizes  $U1$ ,  $U2$  e  $U3$ , e,  $SV1$ ,  $SV2$  e  $SV3$  na classe `uint8` para compararmos a quantidade de bytes necessários para armazená-las com os bytes necessários para armazenar a matriz  $A$ . Temos que a quantidade de bytes necessários é  $3 \times Ui + 3 \times SVi$ , ou seja basta verificarmos quantos bytes são necessários para armazenar  $U18$  e  $SV18$ . Faremos isto digitando as linhas de comando abaixo.

```
U1=(u1(:,1:p)); SV1=(s1(1:p,1:p)*v1(:,1:p)');
U18=(uint8(U1));SV18=(uint8(SV1));
```

Obtemos que para um posto  $p = 50$  necessitamos de 100.050 bytes e para  $p = 100$  de 200.100 bytes, ou seja, aproximadamente 32,5% e 65%, respectivamente, dos bytes que seriam necessários para armazenar a matriz  $A$ . Observando que a razão matricial de compressão, para este caso, é  $\frac{242+425}{242 \times 425} p = \frac{667}{102850} p$ , temos que a razão de compressão para  $p = 10$ ,  $p = 50$  e  $p = 100$  é respectivamente  $\frac{667}{10285} \cong 0.0648$ ,  $\frac{667}{2057} \cong 0.3242$  e  $\frac{1334}{2057} \cong 0.6485$ ; ou seja a mesma taxa que a verificada nos bytes.

## 9 Conclusão

Analisando os vários aspectos que satisfazem os objetivos deste trabalho chegamos as seguintes conclusões:

Refazendo o mesmo exemplo que Dan Kalman (3) publicou na revista *College Mathematics Journal* (1996, 2-23 p). Descobrimos, depois de muito trabalho, que erroneamente as figuras 8 da página 20 e 9 da página 21 publicadas como sendo a aproximação da imagem com  $p=5$ , na realidade é  $p=4$ . Abaixo vemos a imagem original (esquerda) e a imagem publicada como sendo  $p = 5$  (direita).

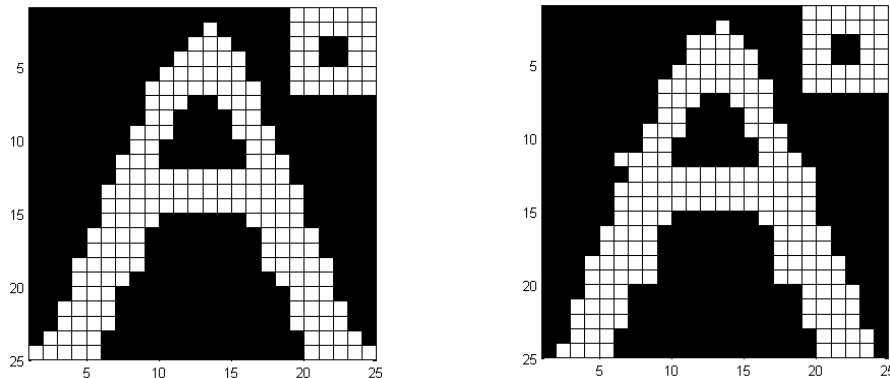


Figura 17: Imagem original (esquerda) e sua compressão com  $p = 4$  (direita)

O presente trabalho contribuiu, de um modo geral, para o melhor entendimento de como os conhecimentos em álgebra linear são importantes às ciências aplicadas, em especial à computação. Os resultados obtidos podem ser compreendidos por qualquer pessoa que saiba um pouco de álgebra linear.

Os experimentos realizados nos permitem dizer que a compressão de imagens digitais via DVS é algo eficiente, pois o uso da DVS resulta em taxas satisfatórias de compressão.

Um software eficiente, como o MatLab, entende e manipula com facilidade a álgebra linear. Portanto, o Matlab<sup>©</sup> contribuiu amplamente na geração, visualização e compressão das imagens.

# Referências

- 1 DEMMEL, James W. *Applied Numerical Linear Algebra*. Society for Industrial and Applied Mathematics - SIAM. Philadelphia, 1997.
- 2 GOLUB, Gene H. & LOAN, Charles F. Van. *Matrix Computations, fourth printing*. Johns Hopkins University Press. Baltimore, 1993
- 3 KALMAN, Dan. A Singularly Valuable Decomposition: The SVD of a Matrix. *College Mathematics Journal* 27 (1996) 2-23.
- 4 LAY, David C. *Álgebra Linear e Suas Aplicações, 2ª edição*. Livros Técnicos e Científicos Editora - LCT. Rio de Janeiro, 1999.
- 5 LEON, Steven J. *Álgebra Linear com Aplicações*. Livros Técnicos e Científicos Editora - LCT. Rio de Janeiro, 1999.
- 6 LIMA, Elon Lages. *Álgebra Linear, 5ª edição*. Instituto de Matemática Pura e Aplicada - IMPA. Rio de Janeiro, 1995.
- 7 LIPSCHUTZ, Seymour. *Álgebra Linear, Série Schaum*. Editora McGraw-Hill do Brasil, LTDA. Brasília, 1973.
- 8 LONG, Andy & LONG, Cliff. Surface Approximation and Interpolation via Matrix SVD. *College Mathematics Journal* 32 (2001) 20-25.
- 9 MULCHY, Colm & ROSSI, John. A Fresh Approach to the Singular Value Decomposition. *College Mathematics Journal* 29 (1998) 199-207.
- 10 NOBLE, Ben & DEMMEL, James W. *Applied Linear Algebra*. Prentice-Hall. New Jersey, 1988.
- 11 STRANG, Gilbert. *Linear Algebra and Its Applications, third edition*. Harcourt Brace & Company Publishers. Massachusetts, 1988.
- 12 ZIZLER, Peter & FRASER, Holly. Eigenpictures and Singular Values of a Matrix. *College Mathematics Journal* 28 (1997) 59-62.