Instituto Superior Técnico 2009/10 – 1º semestre
Álgebra Linear 1º ano da Licenciatura em Engenharia Informática e de Computadores

A matemática do *Google*

No essencial o bem conhecido motor de busca chamado *Google* usa métodos para ordenar os elementos de um conjunto de "páginas" da internet (muitas dezenas de milhar de milhões), cada uma identificada por um termo — palavra, frase ou expressão linguística — contendo "textos" (textos e imagens) com "hiperligações"; por exemplo o termo "mathematics of Google" remete para cerca de uma centena de "páginas" relacionadas entre si e em geral cada termo buscado remete para um conjunto S com algumas centenas ou milhares de "páginas" (identificadas por certos termos sendo por isso que aos elementos de S se chama simplesmente "termos"). A ordenação das "páginas" deve ser feita de acordo com a "importância" de cada uma, do ponto de vista do "público que usa o Google" com todas as reservas atinentes ao contra-aforismo *vox populi vox diaboli*.

Um dos métodos usados permite ordenar para cada termo buscado o correspondente conjunto S, partindo de uma ordenação arbitrária $(t_1, t_2, \ldots, t_n)$ dos elementos de S.
Os elementos de S podem ser figurados como pontos eventualmente ligados por flechas — de modo a obter um *grafo direccionado* Γ — estando essas ligações (*flechas* de Γ) entre esses pontos (*vértices* de Γ) codificadas numa matriz $A = [a_{ij}]$ de tipo $n \times n$, cujos elementos são apenas os números 0 ou 1 — dita matriz das *ligações* de Γ — de acordo com o critério seguinte:

$$a_{ij} = \begin{cases} 1 \text{ se o termo } t_i \text{ figura no texto anexado a } t_j \\ 0 \text{ se o termo } t_i \text{ não figura no texto anexado a } t_j \end{cases}$$

e como nenhum texto se cita a si próprio tem-se $a_{ii} = 0$ para todo o $i = 1, \ldots, n$; sem perda de generalidade pode assumir-se que partindo de um qualquer vértice de Γ pode chegar-se a qualquer outro vértice de Γ através de um caminho constituído por flechas de Γ, dizendo-se então que o grafo Γ é *irredutível*. Note-se que qualquer matriz $A = [a_{ij}]$ de tipo $n \times n$, cujos elementos são apenas os números 0 ou 1 — sem que necessariamente valha a restrição $a_{ii} = 0$ para todo o $i = 1, \ldots, n$ — pode associar-se a um grafo com n vértices e uma tal matriz dir-se-á *irredutível* se o correspondente grafo for irredutível.
Sendo S um conjunto de termos associado a um dado grafo direccionado, com uma matriz de *ligações* $A = [a_{ij}]$ de tipo $n \times n$, a "importância" de cada termo t_j de S mede-se dialecticamente analisando dois aspectos, o primeiro dos quais é:

> 1) Quais os termos em S em que t_j é *citado* (muitos? poucos? e qual a importância dos termos que referem t_j ?)

Há pois que medir a relevância de cada termo de S como *fonte de autoridade*; isso é feito considerando a soma dos elementos de cada coluna

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix} \tag{*}$$

de A. Quer dizer: $c_j = \sum_{i=1}^{n} a_{ij}$ é o número de vezes que o termo t_j é *citado* nos textos anexados aos termos figurando em S:

$$\begin{array}{cc} & \begin{array}{cccc} & t_j & & \\ & \downarrow & & \end{array} \\ t_i \;\rightarrow & \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ a_{21} & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 \end{bmatrix} = A \quad , \\ & \begin{array}{cccc} \downarrow & \downarrow & \cdots & \downarrow \\ c_1 & c_2 & \cdots & c_n \end{array} \end{array}$$

sendo de notar que é afinal essa coluna (*), que é cómodo por razões de economia gráfica escrever como transposta da linha

$$\begin{bmatrix} a_{1j} & a_{2j} & \cdots & a_{nj} \end{bmatrix} \quad ,$$

ou seja que é cómodo escrever como

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix} = \begin{bmatrix} a_{1j} & a_{2j} & \cdots & a_{nj} \end{bmatrix}^t \quad ,$$

que codifica a informação que conduz à medição da importância de t_j como fonte de autoridade.
Se da ordenação dos c_j por ordem de grandeza decrescente resultasse $c_{j_1} > c_{j_2} > \ldots > c_{j_n}$ poderia dizer-se numa primeira aproximação que a ordenação dos elementos de S por ordem decrescente de importância seria: $t_{j_1} \succ t_{j_2} \succ \ldots \succ t_{j_n}$. O n-plo $\boldsymbol{c} = (c_1, c_2, \ldots, c_n)$, que é cómodo escrever como coluna,

$$\boldsymbol{c} = (c_1, c_2, \ldots, c_n) \equiv \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

fornece pois uma medida da importância dos termos de S como *fonte de autoridade* atribuindo a cada termo t_j um *peso de autoridade* c_j.

Porém a “importância” de cada termo t_i de S mede-se também por um outro aspecto:

> 2) Quais os termos em S que t_i *refere* (muitos? poucos? consoante a importância de t_i assim os termos referidos por t_i assumirão maior ou menor importância, sendo mais promovidos ou menos promovidos).

Há pois que medir a relevância de cada termo de S como *fonte de promoção*; isso é feito considerando a soma dos elementos de cada linha

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix} \quad , \tag{**}$$

de A. Quer dizer: $r_i = \sum_{j=1}^n a_{ij}$ é o número de vezes que o texto anexado ao termo t_i *refere* textos anexados a termos figurando em S:

$$\begin{array}{c} \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ a_{21} & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 \end{bmatrix} \\ \begin{matrix} \downarrow & \downarrow & \cdots & \downarrow \\ c_1 & c_2 & \cdots & c_n \end{matrix} \end{array} \begin{array}{cc} \rightarrow & r_1 \\ \rightarrow & r_2 \\ \vdots & \vdots \\ \rightarrow & r_n \\ & \\ & \end{array}$$

sendo de notar que é afinal essa linha (**) que codifica a informação que conduz à medição da importância de t_j como fonte de promoção.
O n-plo $\boldsymbol{r} = (r_1, r_2, \ldots, r_n)$, que é cómodo escrever como coluna,

$$\boldsymbol{r} = (r_1, r_2, \ldots, r_n) \equiv \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}$$

fornece pois uma medida da importância dos termos de S como *fonte de promoção* atribuindo a cada termo t_i um *peso promocional* r_i.
Cada termo $t_k \in S$ é assim afectado de um *peso de autoridade* c_k e de um *peso promocional* r_k.

Seria razoável afectar cada uma das colunas

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix}$$

relativa ao termo t_j de um coeficiente de ponderação — a multiplicar pela coluna — fornecido pelo correspondente peso de autoridade:

$$c_j \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix}$$

de modo que a coluna obtida através de

$$c_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + c_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} + \ldots + c_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix} = \begin{bmatrix} r'_1 \\ r'_2 \\ \vdots \\ r'_n \end{bmatrix}$$

codificaria, *com vantagem* a relevância de cada termo de S como *fonte de promoção*; é pois natural substituir o n-plo $\boldsymbol{r} = (r_1, r_2, \ldots, r_n)$ pelo n-plo *rectificado* pela ponderação efectuada que é o n-plo $\boldsymbol{r}' = (r'_1, r'_2, \ldots, r'_n)$; ora como é fácil ver, a coluna

$$\boldsymbol{r}' = \begin{bmatrix} r'_1 \\ r'_2 \\ \vdots \\ r'_n \end{bmatrix} = \begin{bmatrix} r'_1 & r'_2 & \cdots & r'_n \end{bmatrix}^t$$

nada mais é que a coluna

$$A\boldsymbol{c} = \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ a_{21} & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

ou seja

$$\boldsymbol{r}' = A\boldsymbol{c}\ .$$

Analogamente seria razoável afectar cada uma das linhas

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}$$

relativa ao termo t_i de um coeficiente de ponderação — a multiplicar pela linha — fornecido pelo correspondente peso promocional:

$$r_i \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}$$

de modo que a linha obtida através de

$$r_1\,[a_{11}\ a_{12}\ \cdots\ a_{1n}] + r_2\,[a_{21}\ a_{22}\ \cdots\ a_{2n}] + \ldots + r_n\,[a_{n1}\ a_{n2}\ \cdots\ a_{nn}] = \begin{bmatrix} c'_1 & c'_2 & \cdots & c'_n \end{bmatrix}$$

codificaria, *com vantagem* a relevância de cada termo de S como *fonte de autoridade*; é pois natural substituir o n-plo $\boldsymbol{c} = (c_1, c_2, \ldots, c_n)$ pelo n-plo *rectificado* pela ponderação efectuada que é o n-plo $\boldsymbol{c}' = (c'_1, c'_2, \ldots, c'_n)$; ora como é fácil ver, a coluna

$$\boldsymbol{c}' = \begin{bmatrix} c'_1 \\ c'_2 \\ \vdots \\ c'_n \end{bmatrix} = \begin{bmatrix} c'_1 & c'_2 & \cdots & c'_n \end{bmatrix}^t$$

nada mais é que a coluna

$$A^t \boldsymbol{r}' = \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ a_{21} & 0 & \cdots & a_{21} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 \end{bmatrix}^t \begin{bmatrix} r_1' \\ r_2' \\ \vdots \\ r_n' \end{bmatrix}$$

onde A^t é a transposta da matriz A, ou seja

$$\boldsymbol{c}' = A^t \boldsymbol{r}'$$

e portanto

$$\boldsymbol{c}' = A^t \boldsymbol{r}' = A^t A \boldsymbol{c}\,;$$

a relação

$$\boldsymbol{c}' = A^t A \boldsymbol{c}$$

sugere um processo iterativo,

$$\boldsymbol{c}' = A^t A \boldsymbol{c},\ \ \boldsymbol{c}'' = A^t A \boldsymbol{c}',\ \ \boldsymbol{c}''' = A^t A \boldsymbol{c}'',\ \text{etc}$$

para obter uma sucessão de vectores $\boldsymbol{c}, \boldsymbol{c}', \boldsymbol{c}'', \boldsymbol{c}''', \ldots$ codificando, cada vez mais vantajosamente a relevância de cada termo de S como fonte de autoridade. Escrevendo $\boldsymbol{c}^{(0)}, \boldsymbol{c}^{(1)}, \boldsymbol{c}^{(2)}, \boldsymbol{c}^{(3)}, \ldots, \boldsymbol{c}^{(k)}, \ldots$ em lugar de $\boldsymbol{c}, \boldsymbol{c}', \boldsymbol{c}'', \boldsymbol{c}''', \ldots$ teremos

$$\boldsymbol{c}^{(k)} = (A^t A)^k \boldsymbol{c}^0 \qquad (\text{para } k = 1, 2, 3, \ldots),$$

ou ainda

$$\boldsymbol{c}^{(k)} = A^t A \boldsymbol{c}^{(k-1)} \qquad (\text{para } k = 1, 2, 3, \ldots) \quad ;$$

toda a questão está em saber se — nalgum sentido a precisar — a sucessão dos $\boldsymbol{c}^{(m)}$ "converge" para um certo vector $\boldsymbol{c}^{(\infty)} = \lim_{m\to\infty} \boldsymbol{c}^{(m)}$ não nulo.

Observação.

Que essa convergência nem sempre é assegurada é o que mostra o exemplo obtido com

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

pois neste caso tem-se

$$A^t A = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

e se existisse alguma sucessão $\boldsymbol{c}^{(m)}$ como se definiu atrás — i.e. tal que $\boldsymbol{c}^{(k)} = A^t A \boldsymbol{c}^{(k-1)}$ para $k = 1, 2, 3, \ldots$ — e ela fosse "convergente" para certo vector $\boldsymbol{c}^{(\infty)} = \lim_{m\to\infty} \boldsymbol{c}^{(m)}$ não nulo ter-se-ia

$$\boldsymbol{c}^{(\infty)} = \lim_{m\to\infty} \boldsymbol{c}^{(k)} = A^t A \lim_{m\to\infty} \boldsymbol{c}^{(k-1)} = A^t A \boldsymbol{c}^{(\infty)}$$

e portanto

$$\boldsymbol{c}^{(\infty)} = A^t A \boldsymbol{c}^{(\infty)}$$

ou seja que em particular a matriz $A^t A$ teria um valor próprio igual a 1 e $\boldsymbol{c}^{(\infty)}$ seria um vector próprio associado a 1. Sucede porém que o polinómio característico de $A^t A$ é $p_A(\lambda) = \lambda^3 - 5\lambda^2 + 6\lambda - 1$ de que obviamente 1 não é raiz pelo que 1 não pode ser valor próprio de $A^t A$; quer dizer que não pode haver uma sucessão $\boldsymbol{c}^{(m)}$ como se definiu atrás convergindo para um vector não nulo $\boldsymbol{c}^{(\infty)}$.

Nexte contexto é útil o resultado seguinte, consequência de um teorema conhecido por "teorema de Perron-Frobenius", (Oskar Perron,1880-1975 e Ferdinand Georg Frobenius 1849-1917) que constitui a primeira parte do enunciado seguinte, sendo a segunda parte uma consequência da primeira:

Teorema Sendo $A = [a_{ij}]$ uma matriz de tipo $n \times n$, cujos elementos são apenas os números 0 ou 1 e sendo A irredutível tem-se:

1) Existe um valor próprio $\lambda_{\max}$ de A^tA — com multiplicidade algébrica igual a 1 — e tal que para qualquer outro valor próprio λ de A^tA vale a relação $|\lambda| < \lambda_{\max}$.

2) Considerando em $\mathbb{R}^n$ o produto interno usual e sendo $\boldsymbol{u}^{(0)} \in \mathbb{R}^n$ unitário não ortogonal ao espaço próprio $E_{A^tA}(\lambda_{\max})$ a sucessão definida por

$$\boldsymbol{u}^{(k)} = \frac{(A^tA)^k \boldsymbol{u}^{(0)}}{\|(A^tA)^k \boldsymbol{u}^{(0)}\|} \qquad (\text{para } k = 1, 2, 3, \dots),$$

é uma sucessão convergente — com respeito à norma $\|\cdots\|$ — para um vector próprio $\boldsymbol{u}^{(\infty)}$ pertencente a $E_{A^tA}(\lambda_{\max})$.

Nota.

Note-se que por se ter

$$\boldsymbol{u}^{(k)} = \frac{A^tA\boldsymbol{u}^{(k-1)}}{\|A^tA\boldsymbol{u}^{(k-1)}\|}$$

resulta, tomando limites,

$$\boldsymbol{u}^{(\infty)} = \frac{A^tA\boldsymbol{u}^{(\infty)}}{\|A^tA\boldsymbol{u}^{(\infty)}\|}$$

ou seja

$$A^tA\boldsymbol{u}^{(\infty)} = \|A^tA\boldsymbol{u}^{(\infty)}\|\boldsymbol{u}^{(\infty)}$$

e como $\boldsymbol{u}^{(\infty)} \in E_{A^tA}(\lambda_{\max})$ tem de ser $\lambda_{\max} = \|A^tA\boldsymbol{u}^{(\infty)}\|$.

Exemplo.

Considere-se a matriz de ligações de três termos, $t_1 t_2, t_3$,

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

tendo-se portanto

$$A^tA = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix} \qquad ;$$

o primeiro vector exprimindo a *fonte de autoridade* é

$$\boldsymbol{c} = (2, 1, 2) = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$$

(revelando ambiguidade na autoridade do primeiro e do terceiro termo), de norma $\|\boldsymbol{c}\| = 3$ e os sucessivos vectores unitários exprimindo a *fonte de autoridade* são os vectores $\boldsymbol{u}^{(0)}, \boldsymbol{u}^{(1)}, \boldsymbol{u}^{(2)}, \dots$ calculados adiante, com valores arredondados até à terceira casa decimal e sugerindo o processo de convergência que conduz à seriação $t_1 \succ t_3 \succ t_2$ (embora a ambiguidade se extinga já no cálculo de $\boldsymbol{u}^{(1)}$):

$$\boldsymbol{u}^{(0)} = \frac{\boldsymbol{c}}{\|\boldsymbol{c}\|} = \frac{1}{3}(2, 1, 2) = \frac{1}{3}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} \approx \begin{bmatrix} 0.667 \\ 0.333 \\ 0.667 \end{bmatrix}$$

$$A^tA\boldsymbol{u}^{(0)} = \frac{1}{3}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 7 \\ 3 \\ 6 \end{bmatrix}$$

$$\|A^tA\boldsymbol{u}^{(0)}\| = \frac{1}{3}\sqrt{94}$$

$$\boldsymbol{u}^{(1)} = \frac{A^tA\boldsymbol{u}^{(0)}}{\|A^tA\boldsymbol{u}^{(0)}\|} = \frac{1}{\frac{1}{3}\sqrt{94}}\begin{bmatrix} \frac{7}{3} \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{7}{94}\sqrt{94} \\ \frac{3}{94}\sqrt{94} \\ \frac{3}{47}\sqrt{94} \end{bmatrix} = \frac{1}{\sqrt{94}}\begin{bmatrix} 7 \\ 3 \\ 6 \end{bmatrix} \approx \begin{bmatrix} 0.722 \\ 0.309 \\ 0.619 \end{bmatrix}$$

$$A^tA\boldsymbol{u}^{(1)} = \frac{1}{\sqrt{94}}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}\begin{bmatrix} 7 \\ 3 \\ 6 \end{bmatrix} = \frac{1}{\sqrt{94}}\begin{bmatrix} 23 \\ 10 \\ 19 \end{bmatrix}$$

$$\|A^tA\boldsymbol{u}^{(1)}\| = \frac{3\sqrt{110}}{\sqrt{94}}$$

$$\boldsymbol{u}^{(2)} = \frac{A^tA\boldsymbol{u}^{(1)}}{\|A^tA\boldsymbol{u}^{(1)}\|} = \frac{\sqrt{94}}{3\sqrt{110}}\frac{1}{\sqrt{94}}\begin{bmatrix} 23 \\ 10 \\ 19 \end{bmatrix} = \frac{1}{3\sqrt{110}}\begin{bmatrix} 23 \\ 10 \\ 19 \end{bmatrix} \approx \begin{bmatrix} 0.731 \\ 0.318 \\ 0.604 \end{bmatrix}$$

$$A^tA\boldsymbol{u}^{(2)} = \frac{1}{3\sqrt{110}}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}\begin{bmatrix} 23 \\ 10 \\ 19 \end{bmatrix} = \frac{1}{3\sqrt{110}}\begin{bmatrix} 75 \\ 33 \\ 61 \end{bmatrix}$$

$$\|A^tA\boldsymbol{u}^{(2)}\| = \frac{\sqrt{10\,435}}{3\sqrt{110}}$$

$$\boldsymbol{u}^{(3)} = \frac{A^tA\boldsymbol{u}^{(2)}}{\|A^tA\boldsymbol{u}^{(2)}\|} = \frac{3\sqrt{110}}{\sqrt{10\,435}}\frac{1}{3\sqrt{110}}\begin{bmatrix} 75 \\ 33 \\ 61 \end{bmatrix} = \frac{1}{\sqrt{10\,435}}\begin{bmatrix} 75 \\ 33 \\ 61 \end{bmatrix} \approx \begin{bmatrix} 0.734 \\ 0.323 \\ 0.597 \end{bmatrix}$$

$$A^tA\boldsymbol{u}^{(3)} = \frac{1}{\sqrt{10\,435}}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}\begin{bmatrix} 75 \\ 33 \\ 61 \end{bmatrix} = \frac{1}{\sqrt{10\,435}}\begin{bmatrix} 244 \\ 108 \\ 197 \end{bmatrix}$$

$$\|A^tA\boldsymbol{u}^{(3)}\| = \frac{\sqrt{110\,009}}{\sqrt{10\,435}}$$

$$\boldsymbol{u}^{(4)} = \frac{A^tA\boldsymbol{u}^{(3)}}{\|A^tA\boldsymbol{u}^{(3)}\|} = \frac{\sqrt{10\,435}}{\sqrt{110\,009}}\frac{1}{\sqrt{10\,435}}\begin{bmatrix} 244 \\ 108 \\ 197 \end{bmatrix} = \frac{1}{\sqrt{110\,009}}\begin{bmatrix} 244 \\ 108 \\ 197 \end{bmatrix} \approx \begin{bmatrix} 0.736 \\ 0.326 \\ 0.594 \end{bmatrix}$$

$$A^tA\boldsymbol{u}^{(4)} = \frac{1}{\sqrt{110\,009}}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}\begin{bmatrix} 244 \\ 108 \\ 197 \end{bmatrix} = \frac{1}{\sqrt{110\,009}}\begin{bmatrix} 793 \\ 352 \\ 638 \end{bmatrix}$$

$$\|A^tA\boldsymbol{u}^{(4)}\| = \frac{\sqrt{1159\,797}}{\sqrt{110\,009}}$$

$$\boldsymbol{u}^{(5)} = \frac{A^tA\boldsymbol{u}^{(4)}}{\|A^tA\boldsymbol{u}^{(4)}\|} = \frac{\sqrt{110\,009}}{\sqrt{1159\,797}}\frac{1}{\sqrt{110\,009}}\begin{bmatrix} 793 \\ 352 \\ 638 \end{bmatrix} = \frac{1}{\sqrt{1159\,797}}\begin{bmatrix} 793 \\ 352 \\ 638 \end{bmatrix} \approx \begin{bmatrix} 0.736 \\ 0.327 \\ 0.592 \end{bmatrix}$$

$$A^tA\boldsymbol{u}^{(5)} = \frac{1}{\sqrt{1159\,797}}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}\begin{bmatrix} 793 \\ 352 \\ 638 \end{bmatrix} = \frac{1}{\sqrt{1159\,797}}\begin{bmatrix} 2576 \\ 1145 \\ 2069 \end{bmatrix}$$

$$\|A^tA\boldsymbol{u}^{(5)}\| = \frac{3\sqrt{1358\,618}}{\sqrt{1159\,797}}$$

$$\boldsymbol{u}^{(6)} = \frac{A^tA\boldsymbol{u}^{(5)}}{\|A^tA\boldsymbol{u}^{(5)}\|} = \frac{\sqrt{1159\,797}}{3\sqrt{1358\,618}}\frac{1}{\sqrt{1159\,797}}\begin{bmatrix} 2576 \\ 1145 \\ 2069 \end{bmatrix} = \frac{1}{3\sqrt{1358\,618}}\begin{bmatrix} 2576 \\ 1145 \\ 2069 \end{bmatrix} \approx \begin{bmatrix} 0.737 \\ 0.327 \\ 0.592 \end{bmatrix}$$

$$A^tA\boldsymbol{u}^{(6)} = \frac{1}{3\sqrt{1358\,618}}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}\begin{bmatrix} 2576 \\ 1145 \\ 2069 \end{bmatrix} = \frac{1}{3\sqrt{1358\,618}}\begin{bmatrix} 8366 \\ 3721 \\ 6714 \end{bmatrix}$$

$$\|A^tA\boldsymbol{u}^{(6)}\| = \frac{\sqrt{128\,913\,593}}{3\sqrt{1358\,618}}$$

$$\boldsymbol{u}^{(7)} = \frac{A^tA\boldsymbol{u}^{(6)}}{\|A^tA\boldsymbol{u}^{(6)}\|} = \frac{3\sqrt{1358\,618}}{\sqrt{128\,913\,593}}\frac{1}{3\sqrt{1358\,618}}\begin{bmatrix} 8366 \\ 3721 \\ 6714 \end{bmatrix} = \frac{1}{\sqrt{128\,913\,593}}\begin{bmatrix} 8366 \\ 3721 \\ 6714 \end{bmatrix} \approx \begin{bmatrix} 0.737 \\ 0.328 \\ 0.591 \end{bmatrix}$$

$$A^tA\boldsymbol{u}^{(7)} = \frac{1}{\sqrt{128\,913\,593}}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}\begin{bmatrix} 8366 \\ 3721 \\ 6714 \end{bmatrix} = \frac{1}{\sqrt{128\,913\,593}}\begin{bmatrix} 27167 \\ 12087 \\ 21794 \end{bmatrix}$$

$$\|A^tA\boldsymbol{u}^{(7)}\| = \frac{\sqrt{1359\,119\,894}}{\sqrt{128\,913\,593}}$$

$$\boldsymbol{u}^{(8)} = \frac{A^tA\boldsymbol{u}^{(7)}}{\|A^tA\boldsymbol{u}^{(7)}\|} = \frac{\sqrt{128\,913\,593}}{\sqrt{1359\,119\,894}}\frac{1}{\sqrt{128\,913\,593}}\begin{bmatrix} 27167 \\ 12087 \\ 21794 \end{bmatrix} = \frac{1}{\sqrt{1359\,119\,894}}\begin{bmatrix} 27167 \\ 12087 \\ 21794 \end{bmatrix} \approx \begin{bmatrix} 0.737 \\ 0.328 \\ 0.591 \end{bmatrix}$$

Referências:

1) *The Mathematics of Web Search*, texto acessível na internet em http://www.math.cornell.edu/~mec/Winter2009/RalucaRemus/index.html
2) *The Anatomy of a Large-Scale Hypertextual Web Search Engine*, Sergey Brin e Lawrence Page, texto acessível na internet em http://infolab.stanford.edu/~backrub/google.html
3) *Contemporary Linear Algebra*, Howard Anton e Robert C. Busby, John Wiley, 2003.
4) *Matrix Iterative Analysis*, Richard S. Varga, Springer, 2000.