Numerical Analysis
Problem Sheet 4

Stuart Golodetz

February 14, 2006

1. If $||x|| := \sqrt{x^T x}$ is the usual (Euclidean) length of a vector $x \in \Re^n$, show that the vector Qx has the same length whenever Q is an orthogonal $n \times n$ matrix.

   If we define the angle between vectors $x, y \in \Re^n$ as

$$\angle(x, y) := \cos^{-1}\left(\frac{x^T y}{||x||\ ||y||}\right)$$

   show that the angle between Qx and Qy is unchanged.

Answer

(a) We first note that since Q is an orthogonal matrix, $Q^T = Q^{-1}$. Then:

$$||Qx|| = \sqrt{(Qx)^T Qx} = \sqrt{x^T Q^T Qx} = \sqrt{x^T Q^{-1} Qx} = \sqrt{x^T x} = ||x||$$

(b) We calculate as follows:

$$\angle(Qx, Qy)$$

$$= \cos^{-1}\left(\frac{(Qx)^T Qy}{||Qx||\ ||Qy||}\right)$$

$$= \cos^{-1}\left(\frac{x^T Q^T Qy}{||x||\ ||y||}\right)$$

$$= \cos^{-1}\left(\frac{x^T Q^{-1} Qy}{||x||\ ||y||}\right)$$

$$= \cos^{-1}\left(\frac{x^T y}{||x||\ ||y||}\right)$$

$$= \angle(x, y)$$

2. Show that if $x, y \in \Re^n$ with $x \neq 0$, $y \neq 0$, then the outer product matrix xy^T has an $n-1$-dimensional kernel. Identify the Image (also called the Range) of this matrix. Hence, identify an eigenvector which corresponds to a generally non-zero eigenvalue and give the condition under which this eigenvalue is also zero.

Answer

(a) We consider row-reducing $\mathbf{x}\mathbf{y}^T$. Since $\mathbf{x} \neq 0$ and $\mathbf{y} \neq 0$ there is some leftmost non-zero entry of the matrix at position (i, j) equal to $x_i y_j$. We swap row i with row 1 if necessary, divide the new row 1 by $x_i y_j$ and subtract multiples of the new row 1 from each of the other rows to clear the zeros in the j^{th} column. Since each row is a multiple of $\mathbf{y}$, doing this clears the entire matrix below the first row, leaving us with a matrix of the form:

$$\begin{bmatrix} 0 & \cdots & 0 & 1 & x_i y_{j+1}/x_i y_j & \cdots & x_i y_n / x_i y_j \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

Now if we recall our linear algebra from first year, we'll remember that the rank of the matrix is the number of non-zero rows in its row-reduction. In other words, the size of its image in this case is 1 (there's only one non-zero row). Recalling also the rank-nullity theorem, we get that the sum of the rank and the nullity is n, meaning that the nullity, or the dimension of the kernel, is $n-1$ in this instance, as required.

(b) We want $\mathbf{z}$ s.t. $(\mathbf{x}\mathbf{y}^T)\mathbf{z} = \lambda\mathbf{z}$, i.e. s.t.

$$(\mathbf{x}\mathbf{y}^T)\mathbf{z} = \mathbf{x}(\mathbf{y}^T\mathbf{z}) = \mathbf{x}(\mathbf{y} \bullet \mathbf{z}) = \lambda\mathbf{z}$$

In other words, we want $\mathbf{z}$ s.t.

$$\mathbf{z} = \frac{\mathbf{y} \bullet \mathbf{z}}{\lambda}\mathbf{x} = \mu\mathbf{x} \text{ for some } \mu \in \Re$$

This means $\mathbf{x}$ is an eigenvector of the matrix.

(c) We note that:

$$(\mathbf{x}\mathbf{y}^T)\mathbf{x} = \mathbf{x}(\mathbf{y}^T\mathbf{x}) = (\mathbf{y} \bullet \mathbf{x})\mathbf{x} = \lambda\mathbf{x}$$

So $\lambda = 0$ precisely when $\mathbf{y} \bullet \mathbf{x} = 0$, or when $\mathbf{x}$ is perpendicular to $\mathbf{y}$.

3. If S is a real skew-symmetric matrix, so that $S^T = -S$, and assuming that $I - S$ is non-singular, show that $(I-S)^{-1}(I+S)$ is an orthogonal matrix. (You may want to convince yourself that for a non-singular matrix $(A^{-1})^T = (A^T)^{-1}$.)

Answer

Let $M = (I-S)^{-1}(I+S)$. To show M is orthogonal, we just have to show that its tranpose acts as its inverse. We start by observing that $I - S = (I+S)^T$, since $I - S = I + S^T = I^T + S^T = (I+S)^T$. Calculating the tranpose of M, we get:

$$\begin{aligned}
& [(I-S)^{-1}(I+S)]^T \\
= \; & (I+S)^T((I-S)^{-1})^T \qquad \{\text{since } (AB)^T = B^T A^T\} \\
= \; & (I^T + S^T)((I-S)^T)^{-1} \\
= \; & (I-S)(I^T - S^T)^{-1} \\
= \; & (I-S)(I+S)^{-1}
\end{aligned}$$

Now, multiplying by the original, we have:

$MM^T = (I-S)^{-1}(I+S)(I-S)(I+S)^{-1}$

TODO: I ran out of time to finish this off.

4. By considering the construction via Householder matrices of the QR factorisation of a square and non-singular matrix A, or otherwise, prove that the factors Q and R are unique if the diagonal entries of R are all positive. How many possibilities are there if this restriction is removed?

Answer

We note that we have (as in lectures) $Q^T A = R \Rightarrow Q^T = H(\mathbf{w_{n-1}}) \cdots H(\mathbf{w_2})H(\mathbf{w})$. Now each of these Householder matrices is such that for its given $\mathbf{u}$, $H(\mathbf{w})\mathbf{u} = (\alpha, 0, ..., 0)^T$ where $\alpha = \pm\sqrt{\mathbf{u}^T\mathbf{u}}$. If each diagonal element of R is positive, then we're taking $\alpha = +\sqrt{\mathbf{u}^T\mathbf{u}}$ each time. There's no choice over which to take each time, so there's $1^{(n-1)} = 1$ possibility for R in this case. Since A is non-singular and $Q^T A = R$, we can easily find Q by doing $Q = (RA^{-1})^T$, whence Q is also unique in this case.

If we remove the restriction, there are 2 possibilities for each Householder matrix we use, so there are $2^{(n-1)}$ possibilities overall. (Note that we don't apply a Householder matrix for the last column, so there's only one possibility for that, whence it's not 2^n overall.)

5. By considering the QR factorisation in which the diagonal entries of R are all positive as in the question above (or otherwise), prove that any orthogonal matrix may be expressed as the product of Householder matrices.

Answer

Given any orthogonal matrix A, we observe that AI is a QR factorisation of it, since A is orthogonal and I is upper-triangular. We further observe that there is a (Householder) QR factorisation $A = QR$ with positive elements on the diagonal of R and that this is unique, whence since the diagonal elements of I are all positive as well, $Q = A$ and $R = I$. Now Q is a product of Householder matrices, so A can be expressed as such a product.

6. Determine the eigenvalues of a Householder matrix.

Answer

We pluck from thin air[1] the observations that both $\mathbf{w}$ and $\mathbf{w}^\perp$ are eigenvectors of $H(\mathbf{w})$, since

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

and

$$H(\mathbf{w})\mathbf{w}^\perp = \left(I - \frac{2}{\mathbf{w}^T\mathbf{w}}\mathbf{w}\mathbf{w}^T\right)\mathbf{w}^\perp = \mathbf{w}^\perp - \frac{2}{\mathbf{w}^T\mathbf{w}}\mathbf{w}(\mathbf{w} \bullet \mathbf{w}^\perp) = \mathbf{w}^\perp$$

The eigenvalue corresponding to $\mathbf{w}$ is hence -1 and that corresponding to $\mathbf{w}^\perp$ is 1.

I confess I'm not actually sure how to come up with these analytically[2], but I'd welcome an explanation!

7. TODO

8. For the question, see the problem sheet.

[1] Well, from thin air and random known facts about the eigenvalues of orthogonal matrices...

[2] Though my chances of thinking of a way of doing it would be improved were I not sleep-deprived at present!

Answer

I can't quite get it in the form required, but I can express it as the product of a *lower*-triangular matrix followed by its tranpose (or an upper-triangular matrix *preceded* by its transpose) if that's any good?! We note that:

$$A^T A = (QR)^T QR = R^T Q^T QR = R^T Q^{-1} QR = R^T R$$

In this, R^T is lower-triangular since R was upper-triangular.

9. Show that if $x \in \Re^n$ then

$$J(1,n)J(1,n-1)\cdots J(1,3)J(1,2)x = \begin{bmatrix} \beta \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

for some β and further prove that $\beta^2 = x^T x$.

Answer

(a) The first bit of this is reasonably straightforward. We define:

$$\mathbf{y_k} = J(1,k)\cdots J(1,2)\mathbf{x}$$

Then we observe from lectures that given a vector $\mathbf{z}$ and an index $i \neq 1$, if $\mathbf{w} = J(1,i)\mathbf{z}$ then $w_i = 0$ and $w_j = z_j$ for $j \neq 1$. Well, we can proceed inductively, noting that for some numbers α_k:

$\mathbf{y_2} = J(1,2)\mathbf{x} = [\alpha_2, 0, x_3, x_4, ..., x_n]^T$
$\mathbf{y_3} = J(1,3)\mathbf{y_2} = [\alpha_3, 0, 0, x_4, ..., x_n]^T$
$\vdots$
$\mathbf{y_k} = J(1,k)\mathbf{y_{k-1}} = [\alpha_k, 0, ..., 0, x_{k+1}, ..., x_n]^T$
$\vdots$
$\mathbf{y_n} = J(1,n)\mathbf{y_{n-1}} = [\alpha_n, 0, ..., 0]^T$ where $\beta = \alpha_n$

(b) It remains to show that the value of β^2 is $\mathbf{x}^T\mathbf{x}$. We recall from lectures that $J(i,j)\mathbf{z}$ is really a shorthand for $J(i,j,\theta)\mathbf{z}$, where $\cos\theta = \frac{z_i}{\sqrt{z_i^2+z_j^2}}$ and $\sin\theta = \frac{z_j}{\sqrt{z_i^2+z_j^2}}$. We also recall that if $\mathbf{w} = J(i,j)\mathbf{z}$ then:

$$w_i = \cos\theta \times z_i + \sin\theta \times z_j = \frac{z_i^2}{\sqrt{z_i^2+z_j^2}} + \frac{z_j^2}{\sqrt{z_i^2+z_j^2}} = \sqrt{z_i^2+z_j^2}$$

From this we derive the formula for updating our α values:

$$\alpha_{k+1} = \sqrt{\alpha_k^2 + x_{k+1}^2}$$

What we actually want is the value of β^2, i.e. α_n^2. So we rewrite the above as:

$$\alpha_{k+1}^2 = \alpha_k^2 + x_{k+1}^2$$

We observe that the initial value of the first element, α_1, is just x_1, and compute:

$$\begin{aligned}
&\alpha_2^2 = \alpha_1^2 + x_2^2 = x_1^2 + x_2^2\\
&\alpha_3^2 = \alpha_2^2 + x_3^2 = x_1^2 + x_2^2 + x_3^2\\
&\vdots\\
&\alpha_k^2 = \sum_{i=1}^{k} x_i^2\\
&\vdots\\
&\alpha_n^2 = \beta^2 = \sum_{i=1}^{n} x_i^2 = \mathbf{x}^T\mathbf{x}
\end{aligned}$$

10. Describe a sequence of $(n-1)+(n-2)+\cdots+2+1 = n(n-1)/2$ Givens matrices which when applied sequentially (as premultiplications) will reduce an $n \times n$ matrix to upper triangular form. (Hint: see question above). You have thus established another algorithm for QR factorisation. What is Q?

    Answer

    The required sequence is of the form:

    $$J(n-1,n)(J(n-2,n)J(n-2,n-1))\cdots(J(k,n)\cdots J(k,k+1))\cdots(J(2,n)\cdots J(2,4)J(2,3))(J(1,n)\cdots J(1,3)J(1,2))$$

    Actually we could premultiply the Givens matrices in any order we like to do the reduction, but this is the way which looks similar to that in the previous question. We note that there are clearly the required number of matrices in the above sequence (there's one $J(i,j)$ matrix with $i = n-1$, two with $i = n-2$, ..., $n-1$ with $i = 1$, and $\sum_{i=1}^{n-1} i = \frac{n(n-1)}{2}$). We also note that applying the sequence will reduce a matrix to upper-triangular form since we're systematically zeroing out each element of the matrix below the diagonal.

    So what is Q? Well, we have that $Q^T A = R$, where R is the upper-triangular matrix we've derived by the above method, so Q is the transpose of the above sequence of matrices.

11. If $x \in \Re^n$ show that postmultiplication of the row vector x^T by $J(i,j,\theta)$ with an appropriately chosen value of θ which you should give, will make the j^{th} entry of the resulting row vector equal to zero.

    Answer

    Let $\mathbf{y}^T = \mathbf{x}^T J(i,j,\theta)$. We want to show that with an appropriate value of θ, we can make $y_j = 0$. Well, by transposing we have that:

    $$\mathbf{y} = J(i,j,\theta)^T \mathbf{x}$$

    Now:

$$J(i,j,\theta)^T = \begin{bmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & \cos\theta & & & -\sin\theta & & \\ & & & & 1 & & & & \\ & & & & & \ddots & & & \\ & & & & & 1 & & & \\ & & & \sin\theta & & & \cos\theta & & \\ & & & & & & & 1 & \\ & & & & & & & & \ddots & \\ & & & & & & & & & 1 \end{bmatrix} \begin{matrix} \\ \\ \\ \leftarrow i^{th}\ row \\ \\ \\ \\ \leftarrow j^{th}\ row \\ \\ \\ \\ \end{matrix}$$

$$\uparrow \quad i^{th}\ column \qquad\qquad \uparrow \quad j^{th}\ column$$

So $y_j = \sin\theta \times x_i + \cos\theta \times x_j$, whence we can ensure $y_j = 0$ by choosing $x_i \sin\theta = -x_j \cos\theta$, i.e. $\tan\theta = \frac{-x_j}{x_i}$. Whence $\theta = \tan^{-1}\frac{-x_j}{x_i}$ works to make y_j zero.

12. Using the results of the above questions can you show directly (and constructively) that any matrix $A \in \Re^{n\times n}$ admits a factorisation LQ with L a lower triangular matrix and Q orthogonal. (Note if $B = QR$ then $B^T = R^T Q^T$ certainly is another way of doing this).

Answer

We observe that postmultiplying a matrix A by a Givens matrix $J(i,j,\theta)$, where $\theta = \tan^{-1}\frac{-a_{ij}}{a_{ii}}$, makes the $(i,j)^{th}$ entry of the resulting matrix zero, i.e. $(A \times J(i,j))_{ij} = 0$.

So given matrix A, we postmultiply it by the following sequence of Givens matrices:

$$(J(1,2)J(1,3)\cdots J(1,n))(J(2,3)J(2,4)\cdots J(2,n))\cdots(J(k,k+1)\cdots J(k,n))\cdots(J(n-2,n-1)J(n-2,n))J(n-1,n)$$

This has the effect of making all the entries above the diagonal zero. The result is hence a lower-triangular matrix, which we can call L. Let Q be the transpose of the above sequence. Then Q is certainly orthogonal, since Q^T is the product of a sequence of Givens matrices which are each orthogonal, and the transpose of an orthogonal matrix is also orthogonal. Furthermore, we have that $AQ^T = L$, whence $AQ^{-1} = L$ and $A = LQ$, which satisfies our requirement (L is lower-triangular and Q is orthogonal).

13. For the question, see the problem sheet.

Answer

The output from MATLAB was as follows:

```
EDU>> A = randn(8,8)

A =

   -0.4326    0.3273    1.0668   -0.6918    0.8156   -1.6041   -2.1707    0.3803
   -1.6656    0.1746    0.0593    0.8580    0.7119    0.2573   -0.0592   -1.0091
    0.1253   -0.1867   -0.0956    1.2540    1.2902   -1.0565   -1.0106   -0.0195
```

```
    0.2877    0.7258   -0.8323   -1.5937    0.6686    1.4151    0.6145   -0.0482
   -1.1465   -0.5883    0.2944   -1.4410    1.1908   -0.8051    0.5077    0.0000
    1.1909    2.1832   -1.3362    0.5711   -1.2025    0.5287    1.6924   -0.3179
    1.1892   -0.1364    0.7143   -0.3999   -0.0198    0.2193    0.5913    1.0950
   -0.0376    0.1139    1.6236    0.6900   -0.1567   -0.9219   -0.6436   -1.8740

EDU>> b = ones(8,1)

b =

     1
     1
     1
     1
     1
     1
     1
     1

EDU>> [Q,R] = qr(A)

Q =

   -0.1611   -0.2300    0.5382   -0.3047    0.2537   -0.3863    0.4766   -0.3129
   -0.6204   -0.3852    0.0196    0.2928    0.0786    0.4842   -0.1137   -0.3564
    0.0467    0.1090   -0.0823    0.4536    0.8500   -0.1703   -0.0470    0.1403
    0.1072   -0.2819   -0.1937   -0.5718    0.3713    0.5033    0.1890    0.3431
   -0.4270    0.0614   -0.0381   -0.4716    0.1464   -0.3349   -0.6710    0.0788
    0.4436   -0.7883   -0.0542    0.1050   -0.0391   -0.2405   -0.3219   -0.0695
    0.4429    0.2804    0.3199   -0.1648    0.2008    0.3579   -0.3576   -0.5472
   -0.0140   -0.0594    0.7476    0.1676   -0.0762    0.1895   -0.2008    0.5720


R =

    2.6848    1.0656   -0.7271    0.1488   -1.4897    0.8895    1.2099    0.9289
         0   -2.1697    1.1314   -0.2774    0.5152   -0.5943   -0.8604    0.9816
         0         0    2.2479    0.2619    0.1135   -1.6626   -1.6083   -0.8376
         0         0         0    2.8632   -0.5480   -0.4796   -0.4327   -0.9205
         0         0         0         0    1.8371   -0.7835   -1.0102    0.3578
         0         0         0         0         0    1.6827    0.8039   -0.5434
         0         0         0         0         0         0   -1.8318    0.3748
         0         0         0         0         0         0         0   -1.4277

EDU>> format long
EDU>> x = A\b

x =

   0.11267190627718
   1.17406989720760
   0.92921317139920
   0.19336875271222
   1.27323719973393
  -0.00951437209584
   0.59298582219596
   0.10628256385842

EDU>> y = R\(Q'*b)

y =

   0.11267190627718
   1.17406989720760
   0.92921317139920
   0.19336875271222
   1.27323719973393
  -0.00951437209584
   0.59298582219596
   0.10628256385842
```

The values of $\mathbf{x}$ and $\mathbf{y}$ are clearly identical in the above. This isn't at all surprising, since if $A = QR$ and $A\mathbf{y} = \mathbf{b}$ then $QR\mathbf{y} = \mathbf{b}$ and $R\mathbf{y} = Q^T\mathbf{b}$. Whence $y = R\backslash(Q^T\mathbf{b})$, which is what we asked MATLAB to calculate for us.