Lecture 17 (hrs. 31,33) - November 4, 2025, 15:30 - 18:30 F3

(2.12) Example.

Compute EGP(A) where:

```
A = [ 1, 1, 0, 0;
      2, 2, 1, 0;
     -2, 0, 0,-1;
     -1, 1, 2,-1 ]
```

(*) $A_1 = A$;

(*) $k = 1$; $A_1(1,1) \neq 0 \Rightarrow P_1 = I$; $T_1 = P_1 A_1$;

$$H_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \lambda_2 & 1 & 0 & 0 \\ \lambda_3 & 0 & 1 & 0 \\ \lambda_4 & 0 & 0 & 1 \end{bmatrix}$$

The values $\lambda_2,\lambda_3,\lambda_4$ are determined by requiring that the elements (2,1), (3,1) and (4,1) - that is, the elements of the k-th column below the diagonal - in the matrix $H_1 T_1$ be *equal to zero*:

$$\lambda_2 T_1(1,1) + T_1(2,1) = 0 \quad ; \quad \lambda_3 T_1(1,1) + T_1(3,1) = 0 \quad ; \quad \lambda_4 T_1(1,1) + T_1(4,1) = 0$$

Considering that $T_1(1,1) \neq 0$, the values $\lambda_2,\lambda_3,\lambda_4$ are *uniquely determined*:

$$\lambda_2 = -\frac{T_1(2,1)}{T_1(1,1)} = -2 \quad ; \quad \lambda_3 = -\frac{T_1(3,1)}{T_1(1,1)} = 2 \quad ; \quad \lambda_4 = -\frac{T_1(4,1)}{T_1(1,1)} = 1$$

Finally:

```
[ 1, 0, 0, 0;  [ 1, 1, 0, 0;  =  [ 1, 1, 0, 0;
 -2, 1, 0, 0;    2, 2, 1, 0;       0, 0, 1, 0;
  2, 0, 1, 0;   -2, 0, 0,-1;       0, 2, 0,-1;
  1, 0, 0, 1 ]  -1, 1, 2,-1]       0, 2, 2,-1]

     H1              T1       =        A2
```

(*) $k = 2$; $A_2(2,2) = 0 \Rightarrow$ since $A_2(3,2) \neq 0$, we swap row two and row three: $P_2 = P_{2,3}$;

```
[ 1, 0, 0, 0;  [ 1, 1, 0, 0;  =  [ 1, 1, 0, 0;
  0, 0, 1, 0;    0, 0, 1, 0;       0, 2, 0,-1;
  0, 1, 0, 0;    0, 2, 0,-1;       0, 0, 1, 0;
  0, 0, 0, 1 ]   0, 2, 2,-1]       0, 2, 2,-1]

     P2,3            A2       =        T2
```

so that $T_2(2,2) \neq 0$.

Then:

```
H2 = [ 1, 0, 0, 0;
       0,  1, 0, 0;
       0, λ3, 1, 0;
       0, λ4, 0, 1 ]
```

The values λ_3,λ_4 are determined by requiring that the elements (3,2), and (4,2) - that is, the elements of the k-th column below the diagonal - in the matrix $H_2\,T_2$ be *equal to zero*:

$$\lambda_3 T_2(2,2) + T_2(3,2) = 0 \quad ; \quad \lambda_4 T_2(2,2) + T_2(4,2) = 0$$

Considering that $T_2(2,2) \neq 0$, the values λ_3,λ_4 are *uniquely determined*:

$$\lambda_3 = -\frac{T_2(3,2)}{T_2(2,2)} = 0 \quad ; \quad \lambda_4 = -\frac{T_2(4,2)}{T_2(2,2)} = -1$$

Finally:

```
[ 1, 0, 0, 0;  [ 1, 1, 0, 0;  =  [ 1, 1, 0, 0;
  0, 1, 0, 0;    0, 2, 0,-1;       0, 2, 0,-1;
  0, 0, 1, 0;    0, 0, 1, 0;       0, 0, 1, 0;
  0,-1, 0, 1 ]   0, 2, 2,-1]       0, 0, 2, 0]

     H2              T2       =        A3
```

(*) $k = 3$; $A_3(3,3) \neq 0 \Rightarrow P_3 = I$; $T_3 = A_3$;

```
H3 = [ 1, 0, 0, 0;
       0, 1, 0, 0;
       0, 0, 1, 0;
       0, 0, λ4, 1 ]
```

The value λ_4 is determined by the requirement that the element (4,3) - that is, the elements of the k-th column below the diagonal - in the matrix $H_3\,T_3$ be *equal to zero*:

$$\lambda_4 T_3(3,3) + T_3(4,3) = 0$$

Considering that $T_3(3,3) \neq 0$, the value λ_4 is *uniquely determined*:

$$\lambda_4 = -\frac{T_3(4,3)}{T_3(3,3)} = -2$$

Finally:

```
[ 1, 0, 0, 0;  [ 1, 1, 0, 0;  =  [ 1, 1, 0, 0;
  0, 1, 0, 0;    0, 2, 0,-1;       0, 2, 0,-1;
  0, 0, 1, 0;    0, 0, 1, 0;       0, 0, 1, 0;
  0, 0,-2, 1 ]   0, 0, 2, 0]       0, 0, 0, 0]

     H3              T3       =        A4
```

(*) $D = A_4$; $P = P_3\,P_2\,P_1 = P_{2,3}$;

Then:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}}_{H_1^{-1}} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{P_{2,3}^{t}} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}}_{H_2^{-1}} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{bmatrix}}_{H_3^{-1}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ -2 & 1 & 0 & 0 \\ -1 & 1 & 2 & 1 \end{bmatrix}}_{\Sigma}$$

and then:

$$S = P\,\Sigma = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ -1 & 1 & 2 & 1 \end{bmatrix}$$

The elements $T_1(1,1)$, $T_2(2,2)$ and $T_3(3,3)$ used to derive the elementary Gaussian matrices H_1, H_2 and H_3 (in general, the element $T_k(k,k)$ used to derive the matrix H_k) are called *pivots*. The term *pivoting* refers to the exchanges made at the k-th iteration to obtain $T_k(k,k) \neq 0$.

(2.13) <u>Example</u>.

Compute EGP(A) where:

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

(*) $A_1 = A$;

(*) $k = 1$; $A_1(1,1) \neq 0 \Rightarrow P_1 = I$; $T_1 = P_1\,A_1$;

$$H_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \lambda_2 & 1 & 0 & 0 \\ \lambda_3 & 0 & 1 & 0 \\ \lambda_4 & 0 & 0 & 1 \end{bmatrix}$$

The values $\lambda_2,\lambda_3,\lambda_4$ are determined by requiring that the elements (2,1), (3,1) and (4,1) - that is, the elements of the k-th column below the diagonal - in the matrix $H_1\,T_1$ be *equal to zero*:

$$\lambda_2\,T_1(1,1) + T_1(2,1) = 0 \quad ; \quad \lambda_3\,T_1(1,1) + T_1(3,1) = 0 \quad ; \quad \lambda_4\,T_1(1,1) + T_1(4,1) = 0$$

Considering that $T_1(1,1) \neq 0$, the values $\lambda_2,\lambda_3,\lambda_4$ are *uniquely determined*:

$$\lambda_2 = -\frac{T_1(2,1)}{T_1(1,1)} = -2 \quad ; \quad \lambda_3 = -\frac{T_1(3,1)}{T_1(1,1)} = 2 \quad ; \quad \lambda_4 = -\frac{T_1(4,1)}{T_1(1,1)} = 1$$

Finally:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}}_{H_1} \underbrace{\begin{bmatrix} 1 & 1 & 0 & 0 \\ 2 & 2 & 1 & 0 \\ -2 & 0 & 0 & -1 \\ -1 & 1 & 2 & -1 \end{bmatrix}}_{T_1} = \underbrace{\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 2 & -1 \end{bmatrix}}_{A_2}$$

(*) $k = 2$; $A_2(2,2) = 0 \Rightarrow$ since it is *also* $A_2(3,2) = A_2(4,2) = 0$, the elements of the k-th column below the diagonal are *already equal to zero*, we set: $P_2 = I$ and $H_2 = I$, from which $T_2 = P_2 A_2 = A_2$ and $A_3 = H_2 T_2 = H_2 A_2 = A_2$;

(*) $k = 3$; $A_3(3,3) = 0 \Rightarrow$ since $A_3(4,3) \neq 0$ we swap row three and four: $P_3 = P_{3,4}$, hence:

```
T3 = P3 A3 = [ 1, 1, 0, 0;
               0, 0, 1, 0;
               0, 0, 2,-1;
               0, 0, 0,-1 ]
```

This matrix is already upper triangular, so $H_3 = I$ and $A_4 = T_3$;

(*) $D = A_4$; $P = P_3 P_2 P_1 = P_{3,4}$;

Then:

```
[ 1, 0, 0, 0;  [ 1, 0, 0, 0;   =   [ 1, 0, 0, 0;
  2, 1, 0, 0;    0, 1, 0, 0;         2, 1, 0, 0;
 -2, 0, 1, 0;    0, 0, 0, 1;        -2, 0, 0, 1;
 -1, 0, 0, 1 ]   0, 0, 1, 0 ]       -1, 0, 1, 0 ]
```

H_1^{-1} $\qquad$ $P_{3,4}^t$ $\qquad$ Σ

and:

```
S = P Σ = [ 1, 0, 0, 0;
            2, 1, 0, 0;
           -1, 0, 1, 0;
           -2, 0, 0, 1 ]
```

(2.14) <u>Theorem</u> (existence of LR factorization with pivoting).

Let $A \in R^{n \times n}$. The EGP procedure applied to A returns *a* LR factorization of A with pivoting. That is: *for every* $A \in R^{n \times n}$ there exists *at least one* LR factorization with pivoting.

(<u>Proof</u>: It follows from the two previous examples.)

(2.15) <u>Problem</u> (use of the LR factorization with pivoting).

Let:

```
EGP(A) = ( [ 1, 0, 0;  ,  [ 1, 0, 1;  ,  [ 0, 1, 0;  )  ,  b = [ 1;
             0, 1, 0;       0, 2, 1;       1, 0, 0;               0;
             1, 1, 1 ]      0, 0,-1 ]      0, 0, 1 ]              0 ]
```

Without determining A, decide whether A is invertible and, if so, determine the solution to the system $A x = b$.

(2.16) <u>Procedure</u> (study of a system of linear equations with EGP).

// $A \in R^{n \times n}$, $b \in R^n$.

(S,D,P) = EGP(A);

<u>if</u> there is k such that $d_{kk} = 0$ <u>then</u> STOP; <u>else</u>

c = SA(S,Pb);

$x^* = SI(D,c)$

The procedure is *satisfactory* in the sense that, *however* given the data, it decides whether the matrix is invertible and, if so, determines the solution.

(2.17) <u>Definition</u> (the procedure GS).

A procedure for finding a QR factorization of a matrix $A \in R^{n \times n}$ is the following two-step one, called GS,[1] described in the special case of n = 3.

Let $A = [a_1, a_2, a_3] \in R^{3 \times 3}$.

<u>Step one</u>.

We look for $\Omega = [\omega_1, \omega_2, \omega_3]$ with orthogonal columns and Θ with upper triangular form with $\theta_{kk} = 1$ such that $\Omega\,\Theta = A$. *If such matrices exist*, rewriting the last equality *column by column* we get:

$$\omega_1 = a_1 \quad , \quad \omega_1\,\theta_{1,2} + \omega_2 = a_2 \quad , \quad \omega_1\,\theta_{1,3} + \omega_2\,\theta_{2,3} + \omega_3 = a_3 \qquad (*)$$

The first equality determines ω_1. From the second it follows that:[2]

$$(\omega_1\,\theta_{1,2}) \bullet \omega_1 + \omega_2 \bullet \omega_1 = a_2 \bullet \omega_1$$

Since ω_1 and ω_2 are orthogonal, we have $\omega_2 \bullet \omega_1 = 0$. Then, if $\omega_1 \neq 0$, we *necessarily* have:

$$\theta_{1,2} = (a_2 \bullet \omega_1) \;/\; (\omega_1 \bullet \omega_1)$$

hence:

$$\omega_2 = a_2 - \omega_1\,\theta_{1,2}$$

From the third equality of (*) we then have:

$$(\omega_1\,\theta_{1,3}) \bullet \omega_1 + (\omega_2\,\theta_{2,3}) \bullet \omega_1 + \omega_3 \bullet \omega_1 = a_3 \bullet \omega_1$$

and

$$(\omega_1\,\theta_{1,3}) \bullet \omega_2 + (\omega_2\,\theta_{2,3}) \bullet \omega_2 + \omega_3 \bullet \omega_2 = a_3 \bullet \omega_2$$

Since $\omega_2 \bullet \omega_1 = 0$ and, analogously, $\omega_3 \bullet \omega_1 = 0$, then we *necessarily* have:

$$\theta_{1,3} = (a_3 \bullet \omega_1) \;/\; (\omega_1 \bullet \omega_1)$$

Since it is also $\omega_3 \bullet \omega_2 = 0$, <u>if</u> $\omega_2 \neq 0$, we *necessarily* have:

$$\theta_{2,3} = (a_3 \bullet \omega_2) \;/\; (\omega_2 \bullet \omega_2)$$

and, finally:

$$\omega_3 = a_3 - \omega_1\,\theta_{1,3} - \omega_2\,\theta_{2,3}$$

---

1 The name GS of the procedure derives from that of the *Gram-Schmidt Orthonormalization Procedure*, from which it conceptually derives.

2 Given two columns $v, w \in R^n$, we indicate with $v \bullet w$ their canonical scalar product: $v \bullet w = v_1 w_1 + \ldots + v_n w_n$.

<u>Step two</u>.

The factorization of A found in the previous step is *not* a QR factorization because the columns of Ω do not have unit norm. This second step determines, if possible, a QR factorization by normalizing the columns of Ω.

Let: $\Delta = \mathrm{diag}(\|\omega_1\|,\|\omega_2\|,\|\omega_3\|)$.[3] <u>If</u> it is also $\omega_3 \neq 0$, the matrix Δ it is invertible and it is easily verified that the couple

$$U = \Omega\,\Delta^{-1} \quad , \quad T = \Delta\,\Theta \qquad (**)$$

it is a QR factorization of A. Note that for the upper triangular matrix T, we have:

$$T_{k,k} = \|\omega_k\| > 0$$

(2.18) <u>Theorem</u> (procedure GS and QR factorization).

The GS Procedure described in the previous Definition determines *a* QR factorization of $A \in R^{n\times n}$ if and only if A is invertible.

(<u>Proof</u>. If the procedure does not terminate prematurely because $\omega_k = 0$ for some k, then the pair U,T determined by (**) consists of two invertible matrices (U because it is orthogonal, T because it is triangular with the non-zero norms of the columns ω_k on the diagonal). Conversely, if $\omega_1 = 0$ then $a_1 = 0$ and therefore A is not invertible. If $\omega_1 \neq 0$ and $\omega_2 = 0$ then $0 = a_2 - \omega_1\,\theta_{1,2} = a_2 - a_1\,\theta_{1,2}$, therefore a_1 and a_2 are linearly dependent, therefore A is not invertible. If $\omega_1 \neq 0$, $\omega_2 \neq 0$ and $\omega_3 = 0$...)

(2.19) <u>Remark</u> (non-uniqueness of QR factorization).

Let $A \in R^{n\times n}$ and let U,T be a QR factorization of A. If $E \in R^{n\times n}$ is a diagonal matrix with, for example, E(1,1) = -1 and E(k,k) = 1 for k = 2,...,n, then the pair:

$$U' = U\,E \quad , \quad T' = E\,T$$

is a QR factorization of A *different* from U,T.

(2.20) <u>Procedure</u> (study of a system of linear equations with GS).

// $A \in R^{n\times n}$, $b \in R^n$.

<u>if</u> GS(A) determines $\omega_k = 0$ for some k <u>then</u> STOP; <u>otherwise</u>
    (U,T) = GS(A);
    $x^* = SI(T, U^t\,b)$

This procedure is also *satisfactory* in the sense that, *however* given the data, it decides whether the matrix is invertible (using Theorem (2.18)) and, if so, determines the solution.

3 Borrowing the symbology from *Scilab*, $\mathrm{diag}(v_1,...,v_n)$ indicates the diagonal matrix of size $n \times n$ that has the elements $v_1,...,v_n$ on the main diagonal.

(2.21) Remark (Householder method).

There are procedures that determine a QR factorization of *any* $A \in R^{n \times n}$ (even if non-invertible). For example, the following:

```
(U,T) = Householder(A)

  \\ A ∈ R^(n×n)
  A_1 = A;
  for k = 1,...,n-1 repeat:
       determine an orthogonal matrix X_k ∈ R^(n×n) such that the elements under the main
       diagonal of the first k columns of X_k A_k are zero and set: A_(k+1) = X_k A_k;
  T = A_n;
  U = X_1^t ... X_(n-1)^t
```

Scilab's qr function implements this procedure.

(2.22) Procedure (study of a system of linear equations with Householder).

// $A \in R^{n \times n}$, $b \in R^n$.

(U,T) = Householder(A);
if there is k such that $t_{kk} = 0$ then STOP; otherwise $x^* = SI(T, U^t b)$

This procedure is also *satisfactory*.

(2.1) CONDITIONING OF THE SOLUTION OF A SYSTEM OF LINEAR EQUATIONS

Let:

- $A \in R^{n \times n}$ be an invertible matrix, $b \in R^n$ be a non-zero column and x^* be the solution of the system $A\, x = b$
- $A' \in R^{n \times n}$ be an invertible matrix, $b' \in R^n$ be a column and $\hat{x}$ be the solution of the system $A'\, x = b'$

(2.23) Definition (data perturbations, deviation of the solution).

Let:

$$\delta A = A' - A \in R^{n \times n} \quad , \quad \delta b = b' - b \in R^n$$

be the *data perturbations* and:

$$\delta x = \hat{x} - x^* \in R^n$$

be the *deviation of the solution.*

(2.24) Problema (conditioning of the solution of a system of linear equations).

Given a way to *measure* data perturbations and the deviation of the solution, determine how large the deviation of the solution *can be* as a function of how large the data perturbations *are*.