

Week 5

Householder reflection, QR decomposition. Givens rotation and QR decomposition with Givens rotations. Solution of over-determined systems

QR decomposition

In this topic, the matrices will be matrices with full column rank. This means that if the matrix has m rows and n columns then its rank is $r(A) = n$ (the column vectors are linearly independent).

Our goal is to decompose a matrix $A \in \mathbb{R}^{m \times n}$, $r(A) = n$ in the form $A=QR$, where Q is an $m \times m$ orthogonal matrix (its inverse is its transpose) and R is an $m \times n$ upper triangular matrix.

Later we will see two techniques to perform this decomposition but first let us see how Matlab computes Q and R .

Example 1. (QR decomposition with Matlab)

```
A=[1 2 3
    4 5 6
    7 8 9
    10 11 15]
```

```
A = 4x3
     1     2     3
     4     5     6
     7     8     9
    10    11    15
```

```
[Q,R]=qr(A) % The command is simply qr, the output parameters give the matrices Q and R.
```

```
Q = 4x4
   -0.0776   -0.8331    0.3651   -0.4082
   -0.3105   -0.4512   -0.1826    0.8165
   -0.5433   -0.0694   -0.7303   -0.4082
   -0.7762    0.3124    0.5477   -0.0000
R = 4x3
  -12.8841  -14.5916  -18.6276
         0   -1.0413   -1.1454
         0         0    1.6432
         0         0         0
```

```
% We can see that R is an upper triangular matrix and Q is ortohogonal
% because
```

```
Q'*Q % is the identity matrix
```

```
ans = 4x4
     1.0000     0.0000    -0.0000     0.0000
     0.0000     1.0000         0     0.0000
    -0.0000         0     1.0000    -0.0000
```

```
0.0000    0.0000   -0.0000    1.0000
```

% and the product QR gives back the original matrix A.

Q*R

```
ans = 4x3
    1.0000    2.0000    3.0000
    4.0000    5.0000    6.0000
    7.0000    8.0000    9.0000
   10.0000   11.0000   15.0000
```

The main idea in the construction of the QR decomposition is as follows. We multiply the matrix A by orthogonal matrices, say S_1, S_2, \dots, S_k , from left until we obtain an upper triangular matrix from A. Formally:

$S_k \dots S_2 S_1 A = R$ (R is already an upper triangular matrix). Then we can write

$A = S_1^T S_2^T \dots S_k^T R = QR$, where the matrix $S_1^T S_2^T \dots S_k^T$ is an orthogonal matrix (product of orthogonal matrices) and can be denoted by Q .

Example 2. (QR decomposition with a sequence of orthogonal matrices)

```
A=[1 2 3
    4 5 6
    7 8 9
   10 11 15]
```

```
A = 4x3
     1     2     3
     4     5     6
     7     8     9
    10    11    15
```

% Let us define the following matrices.

```
S1=[-0.077615052570633  -0.310460210282533  -0.543305367994433  -0.776150525706333
     -0.310460210282533   0.910556611158364  -0.156525930472862  -0.223608472104089
     -0.543305367994433  -0.156525930472862   0.726079621672491  -0.391314826182156
     -0.776150525706333  -0.223608472104089  -0.391314826182156   0.440978819739778]
```

```
S1 = 4x4
     -0.0776  -0.3105  -0.5433  -0.7762
     -0.3105   0.9106  -0.1565  -0.2236
     -0.5433  -0.1565   0.7261  -0.3913
     -0.7762  -0.2236  -0.3913   0.4410
```

```
S2=[ 1.0000000000000000    0    0    0
      0 -0.211234790319991  0.350582131484409  0.912399053288811
      0  0.350582131484409  0.898526832371054 -0.264086539969549
      0  0.912399053288811 -0.264086539969549  0.312707957948936]
```

```
S2 = 4x4
     1.0000    0    0    0
     0 -0.2112  0.3506  0.9124
```

```

0    0.3506    0.8985    -0.2641
0    0.9124    -0.2641    0.3127

```

```

S3=[ 1.0000000000000000 0 0 0
      0 1.0000000000000000 0 0
      0 0 -0.997688645856591 0.067951202556251
      0 0 0.067951202556251 0.997688645856592]

```

```

S3 = 4x4
1.0000    0    0    0
0    1.0000    0    0
0    0    -0.9977    0.0680
0    0    0.0680    0.9977

```

% These matrices are orhogonal. Let us check this!

S1*S1'

```

ans = 4x4
1.0000    0.0000    0.0000    -0.0000
0.0000    1.0000    0.0000    -0.0000
0.0000    0.0000    1.0000    -0.0000
-0.0000   -0.0000   -0.0000    1.0000

```

S2*S2'

```

ans = 4x4
1.0000    0    0    0
0    1.0000   -0.0000   -0.0000
0   -0.0000    1.0000   -0.0000
0   -0.0000   -0.0000    1.0000

```

S3*S3' % These are identity matrices indeed.

```

ans = 4x4
1.0000    0    0    0
0    1.0000    0    0
0    0    1.0000    0.0000
0    0    0.0000    1.0000

```

% Construct the product:

R=S3*S2*S1*A % This is an upper triangular matrix that can be denoted by R.

```

R = 4x3
-12.8841   -14.5916   -18.6276
-0.0000   -1.0413   -1.1454
0.0000    0.0000    1.6432
-0.0000    0.0000    0.0000

```

% Thus, matrix Q can be chosen to be

Q=S1'*S2'*S3'

```

Q = 4x4
-0.0776   -0.8331    0.3651   -0.4082

```

-0.3105	-0.4512	-0.1826	0.8165
-0.5433	-0.0694	-0.7303	-0.4082
-0.7762	0.3124	0.5477	-0.0000

% We have obtained the same Q and R matrices as in the previous example.

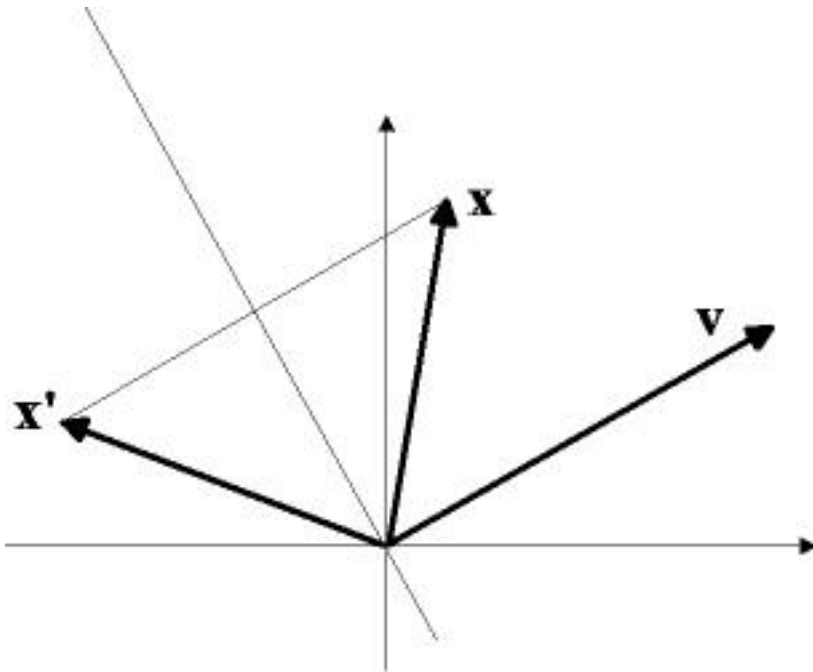
Of course, the question is how to construct the matrices S_1, S_2, \dots that can transform matrix A into an upper triangular form. There are two types of matrices that do this job: Householder reflections and Givens rotations.

Householder reflection

It is easy to see that the matrix $H = I - 2 \frac{vv^T}{v^T v}$ (I is the identity matrix) reflects the vector x through the plane that contains the origin and is orthogonal to the vector v (v is the normal vector of the reflection plane). This fact is based on the form

$$x' = x - \frac{2x^T v}{v^T v} v = x - \frac{2vv^T x}{v^T v} = (I - \frac{2vv^T}{v^T v})x$$

of the reflected vector. This figure demonstrates the reflection in 2D.



Example 3. (reflection through a plane)

% What is the reflection of the vector $x=[1;2;4]$ through the plane with
% normal vector $v=[3;4;5]$?

$x=[1;2;4]$

$x = 3 \times 1$
1
2

```
v=[3;4;5]
```

```
v = 3×1
     3
     4
     5
```

```
H=eye(3)-2*v*v'/(v'*v)
```

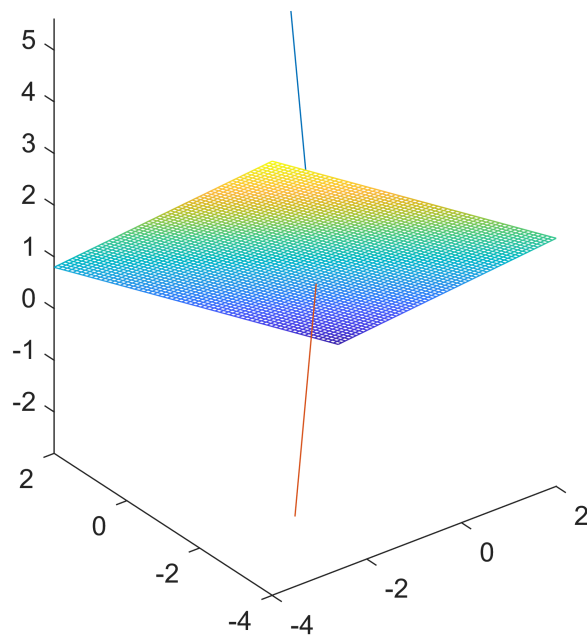
```
H = 3×3
     0.6400    -0.4800    -0.6000
    -0.4800     0.3600    -0.8000
    -0.6000    -0.8000         0
```

```
x_reflected=H*x % This will be the reflected image.
```

```
x_reflected = 3×1
    -2.7200
    -2.9600
    -2.2000
```

```
% Just to see the vectors and the plane in 3D:
```

```
figure
plot3([0,x(1)],[0,x(2)],[0,x(3)])
axis equal
hold on
plot3([0,x_reflected(1)],[0,x_reflected(2)],[0,x_reflected(3)])
hold on
[X,Y]=meshgrid([-4:0.1:2]);
mesh(X,Y,(-v(1)*X-v(2)*Y)/v(3))
hold off
snapnow
```



Two important properties of the reflection matrix H are formulated in Theorem 34.

Thm. 34. H is a symmetric and orthogonal (we need this property to the QR decomposition) matrix.

Proof. The symmetry is trivial. Moreover

$$\left(I - \frac{2vv^T}{v^Tv}\right)\left(I - \frac{2vv^T}{v^Tv}\right) = I - 4\frac{vv^T}{v^Tv} + 4\frac{vv^T}{v^Tv}\frac{vv^T}{v^Tv} = I.$$

Because we would like to use reflection matrices to construct the QR decomposition, it would be nice to choose them (this practically means the choice of the normal vector v) such a way that they zero out as many elements of x as possible. This is achievable. We can zero out all elements of a vector except the first one by choosing the normal vector appropriately. We can reflect the vector x to the first coordinate axis.

Question: How to choose the vector $\bar{\mathbf{v}}$ to reflect the vector $\bar{\mathbf{x}}$ to the axis x_1 , that is parallel to the vector $\bar{\mathbf{e}}_1$?

$$\underbrace{\mathbf{H}\bar{\mathbf{x}}}_{\in \text{lin}(\bar{\mathbf{e}}_1)} = \bar{\mathbf{x}} - \frac{2\bar{\mathbf{v}}^T \bar{\mathbf{x}}}{\bar{\mathbf{v}}^T \bar{\mathbf{v}}} \bar{\mathbf{v}},$$

thus $\bar{\mathbf{v}} \in \text{lin}(\bar{\mathbf{x}}, \bar{\mathbf{e}}_1)$. Let $\bar{\mathbf{v}} = \bar{\mathbf{x}} + \alpha \bar{\mathbf{e}}_1$.

Then

$$\begin{aligned} \mathbf{H}\bar{\mathbf{x}} &= \bar{\mathbf{x}} - \frac{2(\bar{\mathbf{x}}^T + \alpha \bar{\mathbf{e}}_1^T) \bar{\mathbf{x}}}{(\bar{\mathbf{x}} + \alpha \bar{\mathbf{e}}_1)^T (\bar{\mathbf{x}} + \alpha \bar{\mathbf{e}}_1)} (\bar{\mathbf{x}} + \alpha \bar{\mathbf{e}}_1) \\ &= \bar{\mathbf{x}} - 2 \frac{\bar{\mathbf{x}}^T \bar{\mathbf{x}} + \alpha x_1}{\bar{\mathbf{x}}^T \bar{\mathbf{x}} + 2\alpha x_1 + \alpha^2} \bar{\mathbf{x}} - \alpha \frac{2\bar{\mathbf{v}}^T \bar{\mathbf{x}}}{\bar{\mathbf{v}}^T \bar{\mathbf{v}}} \bar{\mathbf{e}}_1 \\ &= \left(1 - 2 \frac{\|\bar{\mathbf{x}}\|_2^2 + \alpha x_1}{\|\bar{\mathbf{x}}\|_2^2 + 2\alpha x_1 + \alpha^2} \right) \bar{\mathbf{x}} - \alpha \frac{2\bar{\mathbf{v}}^T \bar{\mathbf{x}}}{\bar{\mathbf{v}}^T \bar{\mathbf{v}}} \bar{\mathbf{e}}_1. \end{aligned}$$

If $\alpha = \pm \|\bar{\mathbf{x}}\|_2$ then the coefficient of $\bar{\mathbf{x}}$ is zero.

Thus, if a vector $\bar{\mathbf{x}} \neq \mathbf{0}$ is given then $\bar{\mathbf{v}} = \bar{\mathbf{x}} \pm \|\bar{\mathbf{x}}\|_2 \bar{\mathbf{e}}_1$ is a good choice. Then

$$\begin{aligned} \mathbf{H}\bar{\mathbf{x}} &= \mp \|\bar{\mathbf{x}}\|_2 \frac{2(\bar{\mathbf{x}} \pm \|\bar{\mathbf{x}}\|_2 \bar{\mathbf{e}}_1)^T \bar{\mathbf{x}}}{(\bar{\mathbf{x}} \pm \|\bar{\mathbf{x}}\|_2 \bar{\mathbf{e}}_1)^T (\bar{\mathbf{x}} \pm \|\bar{\mathbf{x}}\|_2 \bar{\mathbf{e}}_1)} \bar{\mathbf{e}}_1 \\ &= \mp \|\bar{\mathbf{x}}\|_2 \frac{2\|\bar{\mathbf{x}}\|_2^2 \pm 2\|\bar{\mathbf{x}}\|_2 x_1}{2\|\bar{\mathbf{x}}\|_2^2 \pm 2\|\bar{\mathbf{x}}\|_2 x_1} \bar{\mathbf{e}}_1 = \mp \|\bar{\mathbf{x}}\|_2 \bar{\mathbf{e}}_1. \end{aligned}$$

Def. 35. The reflection matrix \mathbf{H} that reflects a given vector $\bar{\mathbf{x}}$ through a plane that goes through the origin such a way that the reflection is on the first coordinate axis, is called **Householder reflection** (that belongs to the vector $\bar{\mathbf{x}}$).

Application: Based on the above considerations, the Householder reflection that belongs to the vector $\bar{\mathbf{x}}$ can be determined as follows:

- We determine the normal vector of the plane of reflection: $\bar{\mathbf{v}} = \bar{\mathbf{x}} \pm \|\bar{\mathbf{x}}\|_2 \bar{\mathbf{e}}_1$,
- then we construct the reflection matrix with the vector $\bar{\mathbf{v}}$:

$$\mathbf{H} = \mathbf{I} - \frac{2\bar{\mathbf{v}}\bar{\mathbf{v}}^T}{\bar{\mathbf{v}}^T \bar{\mathbf{v}}}.$$

$$\mathbf{H}\bar{\mathbf{x}} = \mathbf{H} \begin{bmatrix} * \\ * \\ \vdots \\ * \end{bmatrix} = \begin{bmatrix} * \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Rmk. If $x_1 \neq 0$ then it is practical to choose the normal vector as $\bar{\mathbf{v}} = \bar{\mathbf{x}} + \text{sgn}(x_1)\|\bar{\mathbf{x}}\|_2\bar{\mathbf{e}}_1$.

Rmk. It is practical to norm the vector $\bar{\mathbf{v}}$ such that the first element of the vector will be 1. Then $\bar{\mathbf{v}}$ can be stored in the place of the eliminated elements of $\bar{\mathbf{x}}$.

Rmk. Let \mathbf{C} be an arbitrary matrix. Then the calculation of $\mathbf{H}\mathbf{C}$ can be performed as follows:

$$\begin{aligned} \mathbf{H}\mathbf{C} &= \left(\mathbf{I} - \frac{2\bar{\mathbf{v}}\bar{\mathbf{v}}^T}{\bar{\mathbf{v}}^T\bar{\mathbf{v}}} \right) \mathbf{C} = \mathbf{C} - \frac{2\bar{\mathbf{v}}\bar{\mathbf{v}}^T}{\bar{\mathbf{v}}^T\bar{\mathbf{v}}} \mathbf{C} \\ &= \mathbf{C} + \underbrace{\bar{\mathbf{v}} \left(-\frac{2\bar{\mathbf{v}}^T \mathbf{C}}{\bar{\mathbf{v}}^T \bar{\mathbf{v}}} \right)}_{=: \bar{\mathbf{w}}^T} = \mathbf{C} + \bar{\mathbf{v}}\bar{\mathbf{w}}^T. \end{aligned}$$

Example 4. (Householder reflection)

```
% Let us give the Householder reflection matrix to the vector x=[-3;4;3;2]
% and perform the transformation.
```

```
x=[-3;4;3;2]
```

```
x = 4x1
    -3
     4
     3
     2
```

```
% We determine the normal vector of the reflection plane
```

```
v=x;
v(1)=v(1)-norm(x) % This will be the normal vector. Here we used subtraction because
```

```
v = 4x1
   -9.1644
    4.0000
    3.0000
    2.0000
```

```
% the first element of x is negative (to avoid catastrophic cancellation)
```

```
% We constuct the Householder reflection
```



```
H=eye(4)-2*v*v'/(v'*v)
```

```
H = 4x4
    -0.4867    0.6489    0.4867    0.3244
    0.6489    0.7168   -0.2124   -0.1416
    0.4867   -0.2124    0.8407   -0.1062
    0.3244   -0.1416   -0.1062    0.9292
```

```
% The result of the transformation
```

```
H*x % Only the first element is nonzero. Its value is the norm of the vector x
```

```
ans = 4x1
    6.1644
    -0.0000
    -0.0000
    -0.0000
```

```
% (reflection does not change the 2-norm of the vectors).
```

With Householder reflection we can construct the QR decomposition of a matrix. This procedure is used also to prove the existence of the QR decomposition of full rank matrices.

Thm. 36. Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ ($m \geq n$) be a full rank matrix. Then there exists an orthogonal matrix $\mathbf{Q} \in \mathbb{R}^{m \times m}$ and an upper triangular matrix $\mathbf{R} \in \mathbb{R}^{m \times n}$ such that $\mathbf{A} = \mathbf{QR}$.

Proof. Let \mathbf{H}_1 be the Householder reflection that belongs to the column $\mathbf{A}(1:m, 1)$. Then the $2:m$ elements of the first column of $\mathbf{A}^{(2)} := \mathbf{H}_1 \mathbf{A}$ are zero. Let $\tilde{\mathbf{H}}_2$ be the Householder reflection that belongs to the column $\mathbf{A}^{(2)}(2:m, 2)$. Moreover, let $\mathbf{H}_2 = \text{diag}(1, \tilde{\mathbf{H}}_2)$. Then the $2:m$ elements of the first column of $\mathbf{A}^{(3)} := \mathbf{H}_2 \mathbf{A}^{(2)}$ and the $3:m$ elements of the second column are zero, etc. Based on the full rank, this procedure can be continued further. We obtain the representation

$$\mathbf{H}_n \cdots \mathbf{H}_1 \cdot \mathbf{A} = \mathbf{R},$$

where \mathbf{R} is an upper triangular matrix. The matrix $\mathbf{Q}^T := \mathbf{H}_n \cdots \mathbf{H}_1$ is orthogonal, so with the above notations we have $\mathbf{A} = \mathbf{QR}$. ■

Example 5. (The QR decomposition of matrix A with Householder reflections from Example 1.)

```
A=[1 2 3
    4 5 6
    7 8 9
    10 11 15]
```

```
A = 4x3
    1    2    3
```

4	5	6
7	8	9
10	11	15

```
% The first Householder matrix to the first column.
```

```
v=A(:,1);
v(1)=v(1)+norm(v);
H1=eye(4)-2*v*v'/(v'*v)
```

```
H1 = 4x4
-0.0776    -0.3105    -0.5433    -0.7762
-0.3105     0.9106    -0.1565    -0.2236
-0.5433    -0.1565     0.7261    -0.3913
-0.7762    -0.2236    -0.3913     0.4410
```

```
% this zeros out the first column of A below the diagonal
```

```
A1=H1*A
```

```
A1 = 4x3
-12.8841   -14.5916   -18.6276
  0.0000     0.2200    -0.2309
  0.0000    -0.3651    -1.9041
  0.0000    -0.9501    -0.5773
```

```
% The second Householder matrix to the second column
```

```
v=A1(2:4,2);
v(1)=v(1)+norm(v);
H2=blkdiag(1,eye(3)-2*v*v'/(v'*v)) % this is a blockdiagonal matrix. The (1,1)
```

```
H2 = 4x4
 1.0000         0         0         0
         0   -0.2112     0.3506     0.9124
         0     0.3506     0.8985    -0.2641
         0     0.9124    -0.2641     0.3127
```

```
% element is 1 and the (2:4,2:4) block is the Householder reflection.
```

```
% this zeros out the second column below the main diagonal
```

```
A2=H2*H1*A
```

```
A2 = 4x3
-12.8841   -14.5916   -18.6276
  0.0000    -1.0413    -1.1454
  0.0000     0.0000    -1.6394
  0.0000    -0.0000     0.1117
```

```
% The third Householder matrix to the third column
```

```
v=A2(3:4,3);
v(1)=v(1)+norm(v);
H3=blkdiag(1,1,eye(2)-2*v*v'/(v'*v))
```

```
H3 = 4x4
    1.0000    0    0    0
    0    1.0000    0    0
    0    0    0.9977 -0.0680
    0    0 -0.0680 -0.9977
```

```
% this zeros out the third column below the main diagonal
```

```
R=H3*H2*H1*A % This is already an upper triangular matrix that can be denoted by R.
```

```
R = 4x3
   -12.8841   -14.5916   -18.6276
    0.0000   -1.0413   -1.1454
    0         0      -1.6432
   -0.0000   -0.0000    0.0000
```

```
% The orthogonal matrix Q can be constructed as follows.
```

```
Q=H1*H2*H3 % Because this matrices are symmetric.
```

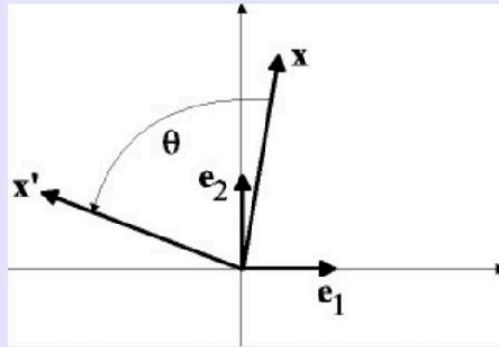
```
Q = 4x4
   -0.0776   -0.8331   -0.3651    0.4082
   -0.3105   -0.4512    0.1826   -0.8165
   -0.5433   -0.0694    0.7303    0.4082
   -0.7762    0.3124   -0.5477    0.0000
```

```
% Compare these Q and R matrices with the result of Example 1. We can see
% that these matrices are very similar to those (some signs are different only)
% and similiary A=QR (the QR decomposition is not unique in general).
```

Givens rotations

There is another possibility to transform a matrix to upper triangular form with orthogonal matrices. We can use rotations instead of reflections. The matrix (orthogonal !) of the rotation with angle θ can be seen below. If the angle θ is chosen such that the second element of the image vector is zero then we say that we have defined a **Given rotation**.

Rotation with angle θ in \mathbb{R}^2 .



$$\bar{\mathbf{x}}' = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \bar{\mathbf{x}}.$$

This matrix is orthogonal. Moreover with the choice $s = \sin \theta$ and $c = \cos \theta$, the vector $[x_1, x_2]^T$ ($x_1 \neq 0$) is transformed to the form $[*, 0]^T$.

- If $x_2 = 0$ then $s = 0$, $c = 1$ is a good choice.
- If $x_2 \neq 0$ then from the solution of the SLAE $sx_1 + cx_2 = 0$, $s^2 + c^2 = 1$ we obtain the parameters

$$s = \frac{\pm x_2}{\sqrt{x_1^2 + x_2^2}}, \quad c = \frac{\mp x_1}{\sqrt{x_1^2 + x_2^2}}.$$

Generally: rotation in the hyperplane (i, j) with angle θ

$$\mathbf{G}(i, j, \theta) = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & c & & -s & \\ & & & 1 & & \\ & & & & \ddots & \\ & s & & & & 1 & c & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{bmatrix}$$

Example 5. (Givens rotation)

```
% Define the Givens rotation matrix to the vector x=[2;5] and to the last
% two elements of the vector y=[1;3;-5;2;5]
```

```
x=[2;5]
```

```
x = 2x1
     2
     5
```

```
y=[1;3;-5;2;5]
```

```
y = 5x1
     1
     3
    -5
     2
     5
```

```
% First we compute the sin and cos values or the rotation angle (we do not need the
% angle explicitly)
```

```
c=x(1)/norm(x), s=-x(2)/norm(x)
```

```
c = 0.3714
s = -0.9285
```

```
% and construct the Givens rotation matrix
```

```
G=[c,-s;s,c]
```

```
G = 2x2
     0.3714     0.9285
    -0.9285     0.3714
```

```
% The effect of G is
```

```
G*x
```

```
ans = 2x1
     5.3852
    -0.0000
```

```
% multiplication with G zeros out the second element of x (the 2-norm does not
% change)
```

```
% For the vector y we can use the same 2D rotation as follows
```

```
blkdiag(1,1,1,G)*y
```

```
ans = 5x1
     1.0000
     3.0000
    -5.0000
     5.3852
    -0.0000
```

```
% this transformation zeros out the last element of the vector y.
```

Because the Givens rotation matrices are orthogonal and they introduce zeros into a vector, they can be used to construct the QR decomposition of a matrix (this is the second method of the production of the QR decomposition). We have to apply the technique of Example 5 in the following way.

QR decomposition (schematically):

$$\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} \rightarrow \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \\ 0 & * & * \end{bmatrix} \rightarrow \begin{bmatrix} * & * & * \\ * & * & * \\ 0 & * & * \\ 0 & * & * \end{bmatrix} \rightarrow \begin{bmatrix} * & * & * \\ 0 & * & * \\ 0 & * & * \\ 0 & * & * \end{bmatrix}$$

$$\begin{bmatrix} * & * & * \\ 0 & * & * \\ 0 & * & * \\ 0 & 0 & * \end{bmatrix} \rightarrow \begin{bmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & * \end{bmatrix} \rightarrow \begin{bmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{bmatrix}$$

Rmk. The number of operations of the Householder QR decomposition is $2n^2(m - n/3)$, while for the Givens QR decomposition we have $3n^2(m - n/3)$.

Example 6. (QR decomposition with Givens rotations)

```
% Let us give the QR decomposition of the matrix A=[1,4;2,6;0,-2;3,7] with
% Givens rotations
```

```
A=[1,4;2,6;0,-2;3,7]
```

```
A = 4x2
```

```
1    4
2    6
0   -2
3    7
```

```
% Givens rotation to the last two elements of the first column.
```

```
x=A(3:4,1);
c=x(1)/norm(x); s=-x(2)/norm(x);
```

```
G1=blkdiag(1,1,[c,-s;s,c])
```

```
G1 = 4x4
```

```
1    0    0    0
0    1    0    0
0    0    0    1
0    0   -1    0
```

```
% The result of the first transformation
```

```
A1=G1*A % the last element in the first column is zero
```

```
A1 = 4x2
     1     4
     2     6
     3     7
     0     2
```

```
% Givens rotation to the 2. and 3. elements of the first column.
```

```
x=A1(2:3,1);
c=x(1)/norm(x); s=-x(2)/norm(x);
```

```
G2=blkdiag(1,[c,-s;s,c],1)
```

```
G2 = 4x4
     1.0000     0     0     0
     0     0.5547     0.8321     0
     0    -0.8321     0.5547     0
     0     0     0     1.0000
```

```
% The result of the second transformation
```

```
A2=G2*G1*A % the last two elements in the first column are zero
```

```
A2 = 4x2
     1.0000     4.0000
     3.6056     9.1526
     0    -1.1094
     0     2.0000
```

```
% Givens rotation to the 1. and 2. elements of the first column.
```

```
x=A2(1:2,1);
c=x(1)/norm(x); s=-x(2)/norm(x);
```

```
G3=blkdiag([c,-s;s,c],1,1)
```

```
G3 = 4x4
     0.2673     0.9636     0     0
    -0.9636     0.2673     0     0
     0     0     1.0000     0
     0     0     0     1.0000
```

```
% The result of the third transformation
```

```
A3=G3*G2*G1*A % the last three elements in the first column are zero
```

```
A3 = 4x2
     3.7417     9.8887
    -0.0000    -1.4084
     0    -1.1094
```

0 2.0000

```
% Now comes the column to zero out
```

```
% Givens rotation to the last two elements of the second column.
```

```
x=A3(3:4,2);  
c=x(1)/norm(x); s=-x(2)/norm(x);
```

```
G4=blkdiag(1,1,[c,-s;s,c])
```

```
G4 = 4x4  
 1.0000    0    0    0  
    0  1.0000    0    0  
    0    0 -0.4851  0.8745  
    0    0 -0.8745 -0.4851
```

```
% The result of the fourth transformation
```

```
A4=G4*G3*G2*G1*A % the last element in the second column is zero, the first
```

```
A4 = 4x2  
 3.7417  9.8887  
-0.0000 -1.4084  
    0  2.2871  
    0    0
```

```
% column does not change
```

```
% Givens rotation to the 2. and 3. elements of the second column.
```

```
x=A4(2:3,2);  
c=x(1)/norm(x); s=-x(2)/norm(x);
```

```
G5=blkdiag(1,[c,-s;s,c],1)
```

```
G5 = 4x4  
 1.0000    0    0    0  
    0 -0.5243  0.8515    0  
    0 -0.8515 -0.5243    0  
    0    0    0  1.0000
```

```
% The result of the fifth transformation
```

```
R=G5*G4*G3*G2*G1*A % this is already an upper triangular matrix, this is why we
```

```
R = 4x2  
 3.7417  9.8887  
    0  2.6859  
 0.0000 -0.0000  
    0    0
```

```
% denoted it by R.
```

```
% The Q matrix can be obtained as follows.
```



```
Q=G1'*G2'*G3'*G4'*G5'
```

```
Q = 4x4
    0.2673    0.5053    0.8205         0
    0.5345    0.2659   -0.3379    0.7276
         0   -0.7446    0.4585    0.4851
    0.8018   -0.3457   -0.0483   -0.4851
```

```
% This is a QR decomposition indeed
```

```
Q*R % gives back A
```

```
ans = 4x2
    1.0000    4.0000
    2.0000    6.0000
    0.0000   -2.0000
    3.0000    7.0000
```

We can see from the number of operations that the CPU time of a QR decomposition for a full matrix with Givens rotations is 50% larger than that of the Householder QR. This is why we generally use Householder's method to produce the QR decomposition. Givens' method is efficient e.g. for upper Hessenberg matrices.

The QR decomposition of an upper Hessenberg matrix (schematically):

$$\begin{bmatrix} * & * & * \\ * & * & * \\ 0 & * & * \\ 0 & 0 & * \end{bmatrix} \rightarrow \begin{bmatrix} * & * & * \\ 0 & * & * \\ 0 & * & * \\ 0 & 0 & * \end{bmatrix} \rightarrow \begin{bmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & * \end{bmatrix} \rightarrow \begin{bmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{bmatrix}$$



Alston Scott Householder, 1904-1993 (USA), Wallace Givens, 1910-1993 (USA)

Over-determined linear systems

Now we consider linear systems with matrices with full column rank. Generally, these are linear systems where we have more equations than unknowns.

$$\mathbf{A}\bar{\mathbf{x}} = \bar{\mathbf{b}}, \quad \mathbf{A} \in \mathbb{R}^{m \times n}, \quad m \geq n, \quad r(\mathbf{A}) = n$$

The above system generally does not have solution (or only one). Then we can search for the vector $\bar{\mathbf{x}}$ (denoted by $\bar{\mathbf{x}}_{LS}$) that minimizes the norm $\|\mathbf{A}\bar{\mathbf{x}} - \bar{\mathbf{b}}\|_2^2$. Let

$$\phi(\bar{\mathbf{x}}) = \|\mathbf{A}\bar{\mathbf{x}} - \bar{\mathbf{b}}\|_2^2,$$

and let $\bar{\mathbf{z}} \in \mathbb{R}^n$ be an arbitrary vector. Because of the full column rank, $\|\mathbf{A}\bar{\mathbf{z}}\|_2 = 0$ can hold only if $\bar{\mathbf{z}} = \mathbf{0}$. Then

$$\begin{aligned} \phi(\bar{\mathbf{x}} + \bar{\mathbf{z}}) &= \|\mathbf{A}(\bar{\mathbf{x}} + \bar{\mathbf{z}}) - \bar{\mathbf{b}}\|_2^2 \\ &= \|\mathbf{A}\bar{\mathbf{x}} - \bar{\mathbf{b}}\|_2^2 + \|\mathbf{A}\bar{\mathbf{z}}\|_2^2 + 2\bar{\mathbf{z}}^T \mathbf{A}^T (\mathbf{A}\bar{\mathbf{x}} - \bar{\mathbf{b}}). \end{aligned}$$

Let $\bar{\mathbf{x}}_{LS}$ be the solution of the SLAE $\mathbf{A}^T \mathbf{A} \bar{\mathbf{x}} = \mathbf{A}^T \bar{\mathbf{b}}$ ($\bar{\mathbf{z}}^T \mathbf{A}^T \mathbf{A} \bar{\mathbf{z}} = \|\mathbf{A}\bar{\mathbf{z}}\|_2^2 \neq 0$ provided that $\bar{\mathbf{z}} \neq \mathbf{0}$, thus $\mathbf{A}^T \mathbf{A}$ is SPD, thus it is non-singular). Then

$$\phi(\bar{\mathbf{x}}_{LS} + \bar{\mathbf{z}}) = \|\mathbf{A}\bar{\mathbf{x}}_{LS} - \bar{\mathbf{b}}\|_2^2 + \|\mathbf{A}\bar{\mathbf{z}}\|_2^2 = \phi(\bar{\mathbf{x}}_{LS}) + \|\mathbf{A}\bar{\mathbf{z}}\|_2^2,$$

that shows that $\bar{\mathbf{x}}_{LS}$ uniquely minimizes ϕ indeed.

Thus the x_{LS} solution (the subscript here means solution in the "least square" sense) minimizes the 2-norm of the residual $r = b - Ax$. The x_{LS} solution can be obtained with two methods: solving the so-called normal equation or using the QR decomposition of the coefficient matrix.

Solution with the normal equation

We have to solve the so-called **normal equation**

$$\mathbf{A}^T \mathbf{A} \bar{\mathbf{x}} = \mathbf{A}^T \bar{\mathbf{b}}.$$

It has unique solution due to the full rank, thus the solution can be written in the form $\bar{\mathbf{x}}_{LS} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \bar{\mathbf{b}}$. This is not efficient in practice.

Computation of $\bar{\mathbf{x}}_{LS}$ with the normal equation

- ▶ $\mathbf{A}^T \mathbf{A}$ is SPD.
- ▶ Let us compute its Cholesky decomposition $\mathbf{L}\mathbf{L}^T$.
- ▶ Let us solve the system $\mathbf{L}\bar{\mathbf{y}} = \mathbf{A}^T \bar{\mathbf{b}}$.
- ▶ We get $\bar{\mathbf{x}}_{LS}$ as the solution of $\mathbf{L}^T \bar{\mathbf{x}} = \bar{\mathbf{y}}$.

Number of operations: $(m + n/3)n^2$ flop

Solution with the QR decomposition

Here we use that multiplications with orthogonal matrices do not change to 2-norm.

Computation of $\bar{\mathbf{x}}_{LS}$ with QR decomposition

$$\begin{aligned} \|\mathbf{A}\bar{\mathbf{x}} - \bar{\mathbf{b}}\|_2^2 &= \|\mathbf{Q}\mathbf{R}\bar{\mathbf{x}} - \bar{\mathbf{b}}\|_2^2 = \|\mathbf{Q}^T(\mathbf{Q}\mathbf{R}\bar{\mathbf{x}} - \bar{\mathbf{b}})\|_2^2 \\ &= \|\mathbf{R}\bar{\mathbf{x}} - \mathbf{Q}^T \bar{\mathbf{b}}\|_2^2 = \|\mathbf{R}_1 \bar{\mathbf{x}} - \bar{\mathbf{c}}\|_2^2 + \|\bar{\mathbf{d}}\|_2^2, \end{aligned}$$

where $\mathbf{R}_1 = \mathbf{R}(1:n, 1:n)$, $\bar{\mathbf{c}} = (\mathbf{Q}^T \bar{\mathbf{b}})(1:n, :)$, $\bar{\mathbf{d}} = (\mathbf{Q}^T \bar{\mathbf{b}})(n+1:m, :)$.

- ▶ Compute the QR decomposition of \mathbf{A} .
- ▶ Determine the matrix $\mathbf{R}_1 = \mathbf{R}(1:n, 1:n)$.
- ▶ Determine the vector $\bar{\mathbf{c}} = (\mathbf{Q}^T \bar{\mathbf{b}})(1:n, :)$.
- ▶ $\bar{\mathbf{x}}_{LS}$ is the solution of the SLAE $\mathbf{R}_1 \bar{\mathbf{x}} = \bar{\mathbf{c}}$.

Number of operations: $2(m - n/3)n^2$ flop

Rmk.

- ▶ If $m \gg n$ then the number of operations of the solution with the QR decomposition is approximately the double of that of the other.
- ▶ For quadratic full rank matrices, the number of operations is the same in both cases: $4n^3/3$, which is the double of that of the Gaussian method. When we take into the account also the memory usage, then the total solution time may be comparable with that of the Gaussian method, moreover, in this case there is no growth factor, that is the method is stable.
- ▶ We cannot use these methods for (nearly) rank deficient matrices.
- ▶ For the normal equation, we can use the CG method but the condition number of the new system will be the square of that of the original system.

Example 7. (Solution of over-determined linear systems)

```
% Solve the over-determined system with A=[1,4;2,6;0,-2;3,7] and b=[1;2;3;4].
```

```
A=[1,4;2,6;0,-2;3,7]
```

```
A = 4x2
```

```
    1    4  
    2    6  
    0   -2  
    3    7
```

```
b=[1;2;3;4]
```

```
b = 4x1
```

```
    1  
    2  
    3  
    4
```

```
% Normal equation
```

```
x=(A'*A)\(A'*b) % This is the x_LS solution
```

```
x = 2x1
```

```
    3.7525  
   -0.9604
```

```
% QR decomposition
```

```
[Q,R]=qr(A) % First we compute the QR decomposition of the matrix A. Here we could use
```

```
Q =
```

```
   -0.2673   -0.5053   -0.0207   -0.8203
```

```

-0.5345    -0.2659     0.7359     0.3194
         0     0.7446     0.4734    -0.4706
-0.8018     0.3457    -0.4837     0.0605
R =
-3.7417    -9.8887
         0    -2.6859
         0         0
         0         0

```

```
% the result of Example 6.
```

```

d=Q'*b;
d=d(1:2)

```

```

d =
-4.5434
 2.5796

```

```
x=R(1:2,1:2)\d % We have to take the upper square part of the matrix R (2x2 block) and
```

```

x =
 3.7525
-0.9604

```

```
% the first 2 element of the vector Q^Tb. This gives x_LS again.
```

```
% Solution with Matlab's built-in function
```

```
% The left division function can solve over-determined systems too.
```

```
x=A\b % This gives x_LS again.
```

```

x =
 3.7525
-0.9604

```

Further problems

Solve problems 38-42. from our problem booklet. Try to solve the problems both in Matlab (using e.g. the previous codes) and manually. In problem 39 use also Householder reflections to obtain the QR decomposition. The solutions can be found below (and also in the m-file of the course) but first give it a try.

Problem 38

```
x=[2;6;-3]; % we define the vector
```

```
% Householder reflection
```

```
v=x; v(1)=v(1)+norm(v) % the normal vector of the reflection plane
```

```

v =
 9

```

6
-3

```
H=eye(3)-2*v*v'/(v'*v) % the Householder matrix
```

```
H =  
-0.2857 -0.8571 0.4286  
-0.8571 0.4286 0.2857  
0.4286 0.2857 0.8571
```

```
H*x % Matrix H zeros out the elements with index greater than 1.
```

```
ans =  
-7.0000  
0.0000  
-0.0000
```

```
% Givens rotation  
c=6/norm(x(2:3)); s=3/norm(x(2:3));  
G=blkdiag(1,[c,-s;s,c]) % The Givens rotation matrix
```

```
G =  
1.0000 0 0  
0 0.8944 -0.4472  
0 0.4472 0.8944
```

```
G*x % Matrix G zeros out the last element of the vector.
```

```
ans =  
2.0000  
6.7082  
0.0000
```

% The manual computations should give the following results (obtained by the symbolic toolbox).

```
x=sym([2;6;-3]); % we define the vector
```

```
% Householder reflection  
v=x; v(1)=v(1)+norm(v) % the normal vector of the reflection plane
```

```
v =  

$$\begin{pmatrix} 9 \\ 6 \\ -3 \end{pmatrix}$$

```

```
H=eye(3)-2*v*v'/(v'*v) % the Householder matrix
```

```
H =
```


$$\begin{pmatrix} -\frac{2}{7} & -\frac{6}{7} & \frac{3}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \end{pmatrix}$$

```
H*x % Matrix H zeros out the elements with index greater than 1.
```

```
ans =
```

$$\begin{pmatrix} -7 \\ 0 \\ 0 \end{pmatrix}$$

```
% Givens rotation
```

```
c=6/norm(x(2:3)); s=3/norm(x(2:3));
```

```
G=blkdiag(1,[c,-s;s,c]) % The Givens rotation matrix
```

```
G =
```

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{2\sqrt{45}}{15} & -\frac{\sqrt{45}}{15} \\ 0 & \frac{\sqrt{45}}{15} & \frac{2\sqrt{45}}{15} \end{pmatrix}$$

```
G*x % Matrix G zeros out the last element of the vector.
```

```
ans =
```

$$\begin{pmatrix} 2 \\ \sqrt{45} \\ 0 \end{pmatrix}$$

Problem 39.

```
A=[0,0;1,3;0,2]
```

```
A =
```

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

```
% Givens rotation
```

```
x=A(1:2,1);
```

```
c=x(1)/norm(x), s=-x(2)/norm(x)
```



```
c = 0
s = -1
```

```
G1=blkdiag([c,-s;s,c],1) % The first Givens matrix
```

```
G1 =
     0     1     0
    -1     0     0
     0     0     1
```

```
A1=G1*A
```

```
A1 =
     1     3
     0     0
     0     2
```

```
x=A1(2:3,2);
c=x(1)/norm(x), s=-x(2)/norm(x)
```

```
c = 0
s = -1
```

```
G2=blkdiag(1,[c,-s;s,c]) % The second Givens matrix
```

```
G2 =
     1     0     0
     0     0     1
     0    -1     0
```

```
R=G2*A1 % R
```

```
R =
     1     3
     0     2
     0     0
```

```
Q=G1'*G2' % Q
```

```
Q =
     0     0     1
     1     0     0
     0     1     0
```

```
Q*R % Check whether A=QR
```

```
ans =
     0     0
     1     3
     0     2
```

```
% Householder reflection
```

```
v=A(:,1);
v(1)=v(1)+norm(v)
```

```
v =
     1
```

```
1
0
```

```
H1=eye(3)-2*v*v'/(v'*v)
```

```
H1 =
    0    -1     0
   -1     0     0
    0     0     1
```

```
% this zeros out the first column of A below the diagonal
```

```
A1=H1*A
```

```
A1 =
   -1   -3
    0    0
    0    2
```

```
% The second Householder matrix to the second column
```

```
v=A1(2:3,2);
v(1)=v(1)+norm(v);
H2=blkdiag(1,eye(2)-2*v*v'/(v'*v))
```

```
H2 =
    1     0     0
    0     0    -1
    0    -1     0
```

```
% this zeros out the second column below the main diagonal
```

```
R=H2*H1*A % This is already R.
```

```
R =
   -1   -3
    0   -2
    0    0
```

```
Q=H1*H2 % This is matrix R.
```

```
Q =
    0     0     1
   -1     0     0
    0    -1     0
```

```
% Built-in function
```

```
[Q,R]=qr(A) % Computation of Q and R with the built-in
```

```
Q =
    0     0     1
   -1     0     0
    0    -1     0
R =
   -1   -3
```

```
0    -2
0     0
```

```
% function, these matrices are the same as we obtained in the previous two parts.
% Here there is no difference between Matlab and manual solutions.
```

Problem 40

Assume that there are two different $A = QR = Q_1 R_1$ decompositions with positive diagonals in R and R_1 . Then we have

$Q^{-1}Q_1 = RR_1^{-1}$ (let us denote this matrix by D).

Here the left hand side is an orthogonal matrix. Its inverse is its transpose. The right hand side is an upper triangular matrix. Its inverse is an upper triangular matrix again. This is possible only if D is a diagonal matrix. At the same time, this matrix must be orthogonal, which means that all its diagonal elements must be $+1$ or -1 .

Thus, $R = DR_1$. Because all diagonal elements in R and in R_1 are positive, D is the identity matrix, which gives the contradiction $R = R_1$.

Problem 41

```
A=[0,0;1,3;0,2] % the matrix from problem 39.
```

```
A =
     0     0
     1     3
     0     2
```

```
b=[3;4;1];
```

```
% built in solver
```

```
A\b
```

```
ans =
    2.5000
    0.5000
```

```
% normal equation
```

```
(A'*A)\(A'*b)
```

```
ans =
    2.5000
    0.5000
```

```
% QR
```

```
[Q,R]=qr(A);
c=Q'*b;
R(1:2,1:2)\c(1:2)
```

```
ans =
    2.5000
    0.5000
```

% Here there is no difference between Matlab and manual solutions.

Problem 42

The value to be minimized can be written in the form

$\left\| \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 3 \\ 2 \end{bmatrix} \right\|_2^2$, which exactly means that we have to compute the x_{LS} solution of the linear system

with matrix $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix}$ and right hand side $\begin{bmatrix} 1 \\ 0 \\ 3 \\ 2 \end{bmatrix}$.

```
A=[1,1,1;
    0,0,1;
    1,1,1;
    4,2,1]
```

```
A =
     1     1     1
     0     0     1
     1     1     1
     4     2     1
```

```
b=[1;0;3;2]
```

```
b =
     1
     0
     3
     2
```

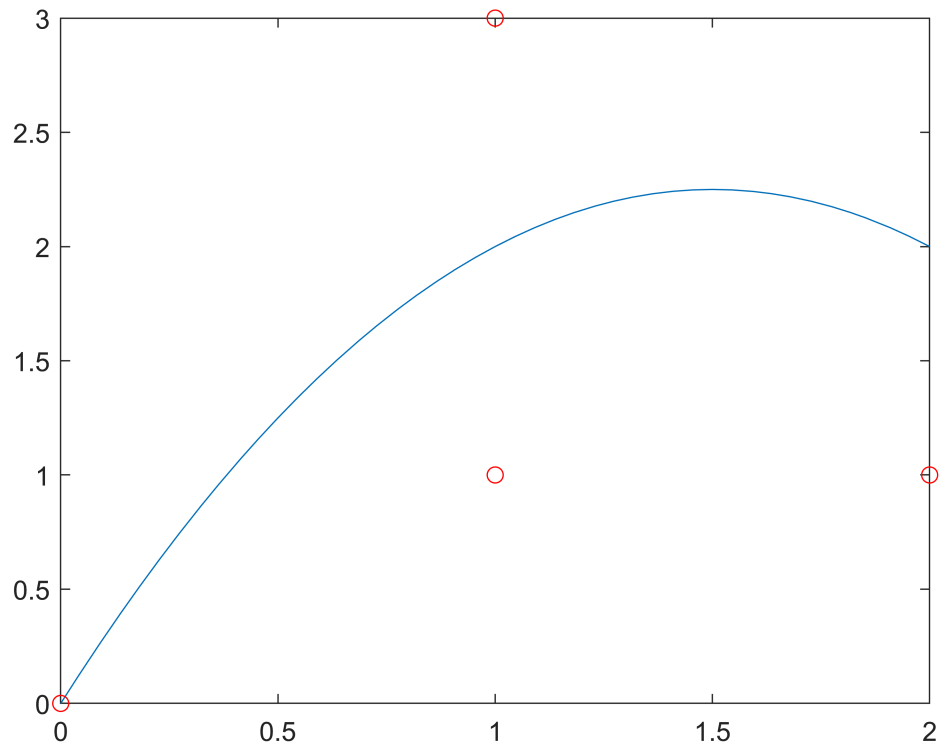
```
x=A\b % We can use here the built-in solver.
```

```
x =
   -1.0000
    3.0000
   -0.0000
```

% Thus the best polynomial is $p(x)=-x^2+3x$

```
% Just to see the points and the fitted polynomial
```

```
figure  
fplot(@(x)-x.^2+3*x,[0,2])  
hold on  
plot([1,0,1,2],[1,0,3,1], 'ro')  
hold off  
snapnow
```



Homework - remember to solve the homework from the webpage! The deadline is 22 March.