TANA15/Lecture 3 - Contents

The QR Decomposition

- Householder Reflections.
- Avoiding the Q.
- Uniqueness.

Structred Least Squares problems

- Givens Rotations and Row updating.

Applications

- Circle fitting, Tikhonov Regularization, Image Deblurring.

Computing the QR Decomposition

Lemma Let $A \in \mathbb{R}^{m\times n}$, $m > n$. Then A can be factorized as

$$A = Q\begin{pmatrix} R \\ 0 \end{pmatrix},$$

where $Q \in \mathbb{R}^{m\times m}$ is *orthogonal* and $R \in \mathbb{R}^{n\times n}$ is triangular.

Remark Typically $m >> n$ so R is small but Q is very large.

Reflections

Definition A *Householder reflection* is a matrix of the form,

$$H(v) = I - 2\frac{vv^T}{v^Tv}.$$

Lemma Let $x \in \mathbb{R}^m$ and $v = x + \text{sign}(x_1)\|x\|_2 e_1$ then,

$$Hx = \pm\|x\|_2 e_1 = \begin{pmatrix} \pm\|x\|_2 \\ 0 \end{pmatrix}.$$

Remarks Computing Hx requires $\mathcal{O}(m)$ operations. Reflections are orthogonal and symmetric.

Example Let $x = (1, -1, 2, -3)^T$. Then

```
>> v=x; v(1)=v(1)+norm(x);

v' =
  4.8730 -1.0000 2.0000 -3.0000

>> y=x-2*(v'*x)/(v'*v)*v

y' =
 -3.8730 -0.0000 0.0000 -0.0000
```

Example Let A be a $m \times 3$ matrix. Pick $x_1 = A(1:m,1)$, $v_1 = x_1 + \|x_1\|_2 e_1$ and compute

$$H(v_1)A = \begin{pmatrix} \|x_1\| & \gamma^T \\ 0 & \widetilde{A}_2 \end{pmatrix} = \begin{pmatrix} + & + & + \\ 0 & + & + \\ 0 & + & + \\ 0 & + & + \end{pmatrix} = A_2.$$

Next let $x_2 = A_2(2:m,2)$ and $v_2 = x_2 + \|x_2\|_2 e_1$. We obtain

$$\begin{pmatrix} 1 & 0 \\ 0 & H(v_2) \end{pmatrix} A_2 = \begin{pmatrix} x & x & x \\ 0 & + & + \\ 0 & 0 & + \\ 0 & 0 & + \end{pmatrix} = A_3.$$

Finally let $x_3 = A_2(3:m,3)$ and $v_3 = x_3 + \|x_3\|_2 e_1$.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & H(v_3) \end{pmatrix} A_3 = \begin{pmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & + \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} R \\ 0 \end{pmatrix}.$$

We have computed a decomposition $A = QR$ where

$$Q = H(v_1)^T H(v_2)^T H(v_3)^T.$$

Lemma Computing R using reflections requires $\mathcal{O}(mn^2)$ operations. An additional $\mathcal{O}(m^2 n)$ operations are needed to also obtain Q.

Lemma Computing Q_1 requires $\mathcal{O}(mn^2)$ additional operations.

Remark The memory needed to store Q and Q_1 is also significant.

Avoiding the Q

Solve min $\|Ax - b\|_2$ with as little work as possible

Proposition Compute the decomposition of the augmented matrix

$$[A,b] = Q\begin{pmatrix} R \\ 0 \end{pmatrix} = Q\begin{pmatrix} R_1 & \gamma \\ 0 & \alpha \\ 0 & 0 \end{pmatrix}, \quad R_1 \in \mathbb{R}^{n\times n}, \gamma \in \mathbb{R}^n.$$

The least squares solution is $x = R_1^{-1}\gamma$. The residual is $|\alpha|$.

Conclusion We do not need Q explicitly. Solving a least squares problem of size $m \times n$ requires $\mathcal{O}(mn^2)$ operations.

Fitting a circle

Example We are given two vectors x and y containing $m = 13$ points located on, or near, the circle. How to estimate the radius and center?

Model Any point $(x, y)^T$ on a circle satisfies the equation

$$F(u) = a(x^2 + y^2) + b_1 x + b_2 y + 1 = 0,$$

where $u = (a, b_1, b_2)^T$ are unknown parameters.

The center and radius of the circle are

$$z = (x_0, y_0)^T = -(\frac{b_1}{2a}, \frac{b_2}{2a})^T, \quad \text{and,} \quad r^2 = \frac{b_1^2 + b_2^2}{4a^2} - \frac{1}{a}.$$

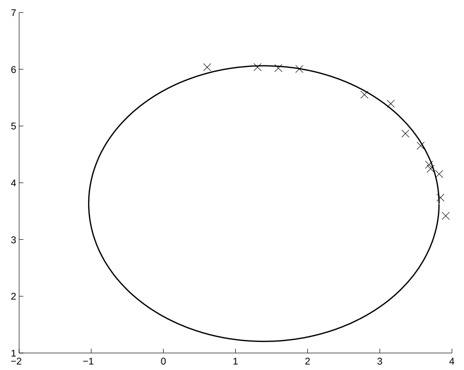
What is the corresponding linear system $Ax \approx b$?

Data points stored in two vectors `x` and `y`. In Matlab

```
>> A=[x.^2+y.^2 x y]; b=-ones(size(x));
>> R=triu( qr( [ A , b ],0) );
>> u=R(1:3,1:3)\R(1:3,4);
>> z=-u(2:3)/2/u(1)
>> r=sqrt( (u(2)^2+u(3)^2)/4/u(1)^2-1/u(1))
```

The result is $z \approx (1.4033, 3.6658)^T$ and $r \approx 2.4215$.

The exact circle was $z = (1.2, 3.4)^T$ and $r = 2.7$. Need more data $(x_k, y_k)^T$ to improve accuracy.



The $m = 13$ data points $(x_k, y_k)^T$ and the estimated circle.

Question How should we proceed to improve the accuracy?

Structured least squares problems

Have computed $A_1 = Q_1 R_1, A_1 \in \mathbb{R}^{13\times 3}$, and minimzed $\|A_1 u - b_1\|_2$.

Given an additional measurement $(x_{14}, y_{14})^T$ we want to minimize

$$\left\| \begin{pmatrix} A_1 \\ a_{14}^T \end{pmatrix} u - \begin{pmatrix} b_1 \\ b_{14} \end{pmatrix} \right\|_2 = \left\| \begin{pmatrix} R_1 \\ a_{14}^T \end{pmatrix} u - \begin{pmatrix} Q_1^T b_1 \\ b_{14} \end{pmatrix} \right\|_2$$

We need the QR decomposition of the *structured matrix*

$$\begin{pmatrix} R_1 & Q_1^T b_1 \\ a_{14}^T & b_{14} \end{pmatrix} = \begin{pmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \\ x & x & x & x \end{pmatrix}$$

Already "almost" triangular! Want to take advantage of the structure!

Givens Rotations

Lemma Suppose $x = (x_1, x_2)^T$ and set

$$\cos(\theta) = \frac{x_1}{\sqrt{x_1^2 + x_2^2}}, \quad \text{and,} \quad \sin(\theta) = \frac{x_2}{\sqrt{x_1^2 + x_2^2}}.$$

Then

$$Gx = \frac{1}{\|x\|_2} \begin{pmatrix} x_1 & x_2 \\ -x_2 & x_1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \|x\|_2 \\ 0 \end{pmatrix}.$$

Definition A *Givens rotation* G_{ij} rotates rows i and j.

Example Let A be a 4×3 matrix. Apply a sequence of Givens Rotations

$$G_{14}G_{13}G_{12} \begin{pmatrix} x & x & x \\ x & x & x \\ x & x & x \\ x & x & x \end{pmatrix} = \begin{pmatrix} + & + & + \\ 0 & + & + \\ 0 & + & + \\ 0 & + & + \end{pmatrix}$$

Next use a_{22} to zero the elements $A(3:4, 2)$, and finally the element a_{33} to zero a_{43}

$$G_{34}G_{24}G_{23} \begin{pmatrix} x & x & x \\ 0 & x & x \\ 0 & x & x \\ 0 & x & x \end{pmatrix} = G_{34} \begin{pmatrix} x & x & x \\ 0 & + & + \\ 0 & 0 & + \\ 0 & 0 & + \end{pmatrix} = \begin{pmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & + \\ 0 & 0 & 0 \end{pmatrix}$$

Thus $A = QR$ where $Q = G_{12}^T G_{13}^T G_{14}^T G_{23}^T G_{24}^T G_{34}^T$.

Observation A matrix $A \in \mathbb{R}^{m \times n}$ has $mn - \frac{1}{2}n^2$ subdiagonal elements. One Givens rotation is needed to zero out each element.

Lemma Computing the R using rotations requires $\mathcal{O}(mn^2)$ operations.

Remark Computing R, Q, or Q_1 using either reflections or rotations requires the same amount of work.

Row updating

Have $A_1 = Q_1 R_1$, where $A_1 \in \mathbb{R}^{m \times n}$ and $R_1 \in \mathbb{R}^{n \times n}$. Want the QR decomposition of

$$\begin{pmatrix} R_1 & Q_1^T b_1 \\ a_{m+1}^T & b_{m+1} \end{pmatrix} = \begin{pmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \\ x & x & x & x \end{pmatrix}$$

First apply a Givens rotation G_{14} and obtain

$$G_{14} \begin{pmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \\ x & x & x & x \end{pmatrix} = \begin{pmatrix} + & + & + & + \\ 0 & x & x & x \\ 0 & 0 & x & x \\ 0 & + & + & + \end{pmatrix}.$$

Continue and apply rotations G_{24} and G_{34} to obtain

$$G_{34}G_{24}\begin{pmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \\ 0 & x & x & x \end{pmatrix} = G_{34}\begin{pmatrix} x & x & x & x \\ 0 & + & + & + \\ 0 & 0 & x & x \\ 0 & 0 & + & + \end{pmatrix} = \begin{pmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & + & + \\ 0 & 0 & 0 & + \end{pmatrix}.$$

Denote the matrix by R_2. The updated least squares solution is

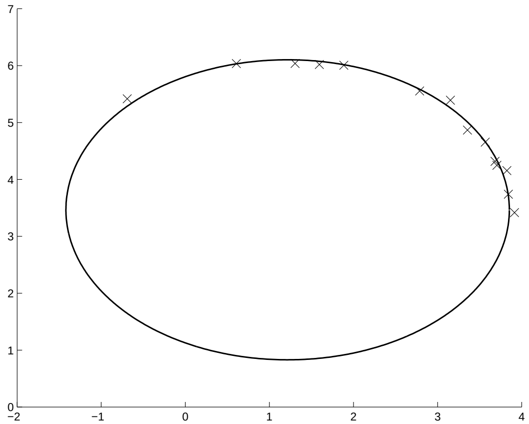
$$x = R_2(1:n,1:n)^{-1}R_2(1:n,n+1).$$

Remark Need n Givens rotations to restore triangular shape! Only $\mathcal{O}(n^2)$ operations. Note that Q is not needed!

Fitting a circle

Example We obtain another measurement (x_{14}, y_{14}) and compute a new row $a_{14}u = b_{14}$. In Matlab:

```
>> a14=[x14.^2+y14.^2 x14 y14]; b14=-1;
>> R = [ R ; a14 , b14 ]
   R =
     -126.7605   -9.9725  -18.1245    3.5816
             0   -4.8810    1.4620    0.3408
             0         0    0.3085   -0.1882
       28.8479   -0.6628    5.3300   -1.0000
 >> for k=1:3, R=Givens( R , k , 4 ); end, R
    R =
      130.0017    9.5768   18.8554   -3.7142
             0    5.6568   -1.8555   -0.2029
             0         0   -0.4133    0.3587
             0         0         0   -0.1195
```

The $m = 14$ data points $(x_k, y_k)^T$ and estimated circle. Now we obtain $z \approx (1.2630, 3.44)^T$ and $r \approx 2.6367$.

Application: Tikhonov Regularization

If a linear system $Ax = b$ is very ill-conditioned need to stabilize the numerical computations.

Definition The *Tikhonov functional* is,

$$f(\lambda) = \min_{x\in\mathbb{R}^n} \|Ax - b\|_2^2 + \lambda^2\|x\|_2^2.$$

Remark Tikhonov regularization is a very popular method for finding approximate solutions to very ill-conditioned systems. Need to solve the problem for many different values of the parameter λ.

This is a structured least squares problem

$$\begin{pmatrix} A \\ \lambda I \end{pmatrix} = \begin{pmatrix} x & x & x \\ x & x & x \\ x & x & x \\ x & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x \end{pmatrix} = \widetilde{Q}^T \begin{pmatrix} x & x & x \\ 0 & x & x \\ 0 & 0 & x \\ x & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x \end{pmatrix}$$

Lemma Suppose we have the decomposition $A = QR$. Then the Tikhonov functional can be evaluated using $\mathcal{O}(n)$ additional Givens rotations.

Proof See L. Eldén, *Algorithms for the regularization of ill–conditioned least squares problems*, BIT, 1977.

Application: Image Deblurring

Example In Astronomy light from stars is collected during several hours. During this time the object moves across the sky. Also light is diffused when it passes through the atmosphere.

Question Can we reduce the influence of these effects using Linear algebra? How to construct a mathematical model?

The Point Spread function

Definition An image consists of a set of pixels $\{I_{ij}\}$ each with a value corresponding to a specific color. Hence the image is a matrix I.

Definition If the exact image is $I_{i_0,j_0} = 1$ and $I_{i,j} = 0$ for $(i,j) \neq (i_0,j_0)$. Then the camera records

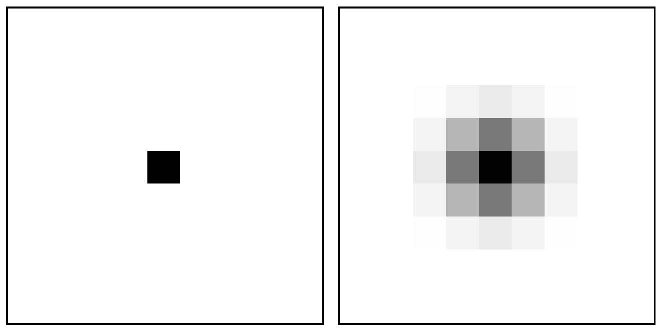
$$I_{ij} = P(i - i_0, j - j_0),$$

where P is the *point spread function*.

Example The blurring effect of light passing through the atmosphere can be modelled as,

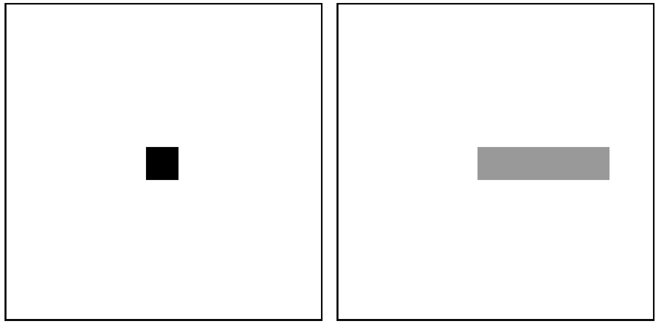
$$P_b(i,j) = \exp\left(\frac{-i^2 - j^2}{2\sigma^2}\right).$$

This means the recorded image is $I_b = A_b I$, where A_b is a matrix.



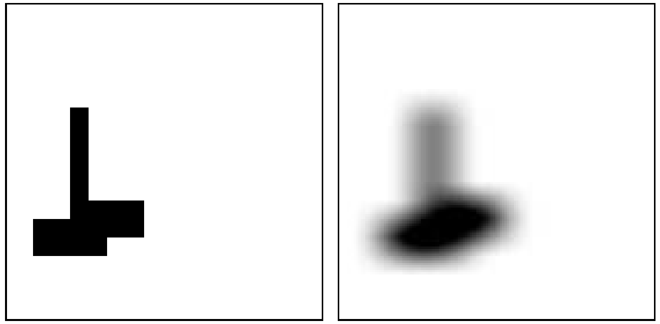
The original 9×9 image I the resulting blurred image $I_b = A_b I$.

Example The effect of object movement during camera exposure. The recorded image is $I_m = A_m I$, where A_m is a matrix.

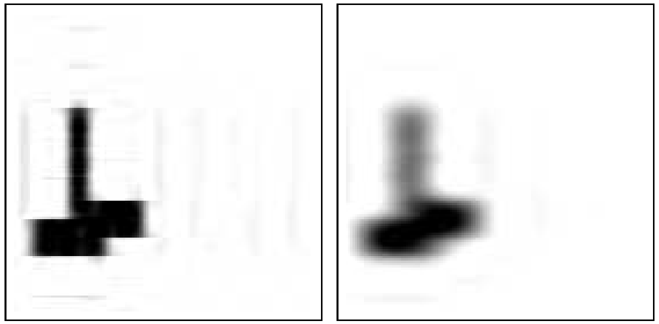


The original 9×9 image I and the resulting image I_m due to object movement. The object appears at several pixels.

Example We look at a sattelite through a telescope. The sattelite moves horizontally through the image. The image we record is I_r which is subject to a combination of object movement and blurring.



The original 128×128 image I showing the "exact" sattelite. Also the degraded image I_r. Can we compensate for these effects?



Results The restored images using Tikhonov regularization. We have $\lambda = 10^{-5}$ (left) and $\lambda = 10^{-2}$ (right). Too much regularization means the image is not clear. The movement is easily compensated for. The blurring is more difficult.

Observation The Matrices A_b and A_m are both 16384×16384 and sparse. It is not feasible to compute the full QR decomposition of the combined matrix $A_T = A_b A_m$.

Alternative techniques are needed for realistic applications.

Uniqueness

Question Is the QR decomposition unique?

Lemma Let R_1, R_2 be upper triangular. Then $R_1 R_2$ and R_1^{-1} are both upper triangular.

Lemma If A is both *triangular* and *orthogonal* then A is diagonal, and $a_{ii} = \pm 1$.

Proof See problem collection.

Lemma Let $A = Q_1 R_1 = Q_2 R_2$ be two different reduced QR decompositions. There exists a diagonal matrix D, $d_{ii} = \pm 1$, and $R_1 = D R_2$.

Remark The QR decomposition is essentially unique. Possibly the diagonal elements of R can change signs.

If $Q = (Q_1, Q_2)$, where Q_1 is the orthogonal basis for $\text{Range}(A)$, then only Q_1 is essentially unique.