TANA09 Lecture 5

Linear systems of equations

- Matrix norms. The condition number. Error analysis.

Linear least squares problems.

- Normal equations.
- Orthogonal matrices. The *QR* decomposition.
- Application - Fit a circle to given points.

Matrix norms

Definition Let $\|\cdot\|$ be a vector norm. A *matrix norm* is given by,

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

Properties $\|A\| \geq 0$, $\|A\| = 0$ if and only if $A = 0$, $\|\alpha A\| = |\alpha|\|A\|$, and $\|A + B\| \leq \|A\| + \|B\|$.

Lemma Let A, B be matrices. Then $\|AB\| \leq \|A\|\|B\|$.

Remark The matrix norm is a measure of how much the *linear operator A* can change a vector? How to compute $\|A\|$?

Lemma For the maximum norm we have

$$\|A\|_\infty = \max_{1 \leq i \leq n} \Big(\sum_{j=1}^{n} |a_{ij}|\Big).$$

It is easy to compute $\|A\|_\infty$. Similar result for $\|A\|_1$.

Lemma For the Euklidean norm we have that

$$\|A\|_2 = \left(\max_{1 \leq i \leq n} \lambda_i(A^T A) \right)^{1/2}.$$

It requires more work to compute $\|A\|_2$. Simpler definition using the *Singular value decomposition*.

Sensitivity analysis

Lemma Let $Ax = b$. If we instead have $b_\delta = b + \delta b$ the resulting error in the solution is bounded by

$$\frac{\|\delta x\|}{\|x\|} \leq \|A\|\|A^{-1}\| \frac{\|\delta b\|}{\|b\|}.$$

Definition The *condition number* is $\kappa(A) = \|A\|\|A^{-1}\|$.

The condition number is a measure of how sentitive a linear system is with respect to changes in the right hand side.

Example Let

$$A = \begin{pmatrix} 4 & -2 & 1 \\ 2 & 1 & 0 \\ -1 & 3 & 1 \end{pmatrix}, \quad \bar{b} = \begin{pmatrix} 6 \\ 2 \\ 0 \end{pmatrix}.$$

Find an upper bound for the error in x measured in $\|\cdot\|_\infty$ if $b_1 = 6 \pm 0.05$, $b_2 = 2 \pm 0.03$, and b_3 is considered exact.

Remark Hard to compute $\kappa(A)$ since A^{-1} is required. Cam be estimated if the decomposition $PA = LU$ is known. In Matlab we have `condest()` and `cond`.

Many problems are *ill-conditioned* with $\kappa(A) \approx 10^{10} - 10^{15}$. Totally different methods are needed.

Example Suppose $\bar{x}$ is an approximate solution to $Ax = b$. How to check the accuracy?

Definiton Given an approximate solution $\bar{x}$ we define a *residual* $r = b - A\bar{x}$.

Lemma The error in $\bar{x}$ can be estimated

$$\frac{\|x - \bar{x}\|}{\|\bar{x}\|} \leq \kappa(A) \frac{\|r\|}{\|A\|\|\bar{x}\|}.$$

Remark If the problem is well-conditioned and the residual r is small then the solution is accurate.

The least squares problem

Example A physical quantity is given by

$$y = Ax_1 + Bx_2,$$

where A and B are constants to be determined and x_1 and x_2 are parameters we can change. How to find A and B?

Solution Perform experiments and measure y for different values of (x_1, x_2).

x_1	x_2	y
3.0	1.60	12.6
2.7	1.35	11.2
1.8	0.70	6.3
4.2	1.20	13.2

$$\Longrightarrow \begin{pmatrix} 3.0 & 1.60 \\ 2.7 & 1.35 \\ 1.8 & 0.70 \\ 4.2 & 1.20 \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 12.6 \\ 11.2 \\ 6.3 \\ 13.2 \end{pmatrix}$$

If we use two measurements we obtain a *linear system*. More measurements should give a better result.

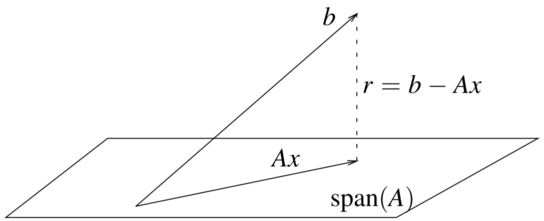
Definition Let $A \in \mathbb{R}^{m \times n}$ $m > n$. The system of equations $Ax = b$ is *over determined*.

More equations than unknowns means a solution usually does not exist. Instead we minimize the *residual* and solve

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|.$$

The case $\|\cdot\|_2$ can be solved easily. This is the *linear least squares problem*.

For minimization in $\|\cdot\|_\infty$ or $\|\cdot\|_1$ se courses in non-linear optimization.



Definiton The vector x that minimize $\|Ax - b\|_2$ is called the *least squares solution*.

Theorem The least squares solution x satisfies the *normal equations*, i.e. $A^T Ax = A^T b$.

If the columns in A are linearly independent then $A^T A$ is non-singular

Example With A and b as before we obtain

$$A^T A = \begin{pmatrix} 37.17 & 14.75 \\ 14.75 & 6.31 \end{pmatrix}, \qquad A^T b = \begin{pmatrix} 134.82 \\ 55.53 \end{pmatrix}.$$

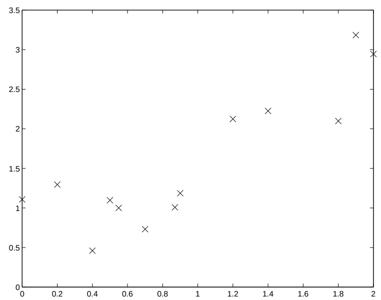
Solve $A^T Ax = A^T b$ gives the *least squares solutuion* $x = (\,1.873\,,\;4.420\,)^T$.

We had picked $x = (\,2\,,\;4\,)^T$ and added a random error to b.

Remark The error estimate for linear systems gives an error bound for $x = (\,A\,,\;B\,)^T$. Usually not very relevant. The condition number $\kappa(A^T A) \approx \kappa(A)^2$.

Model fitting

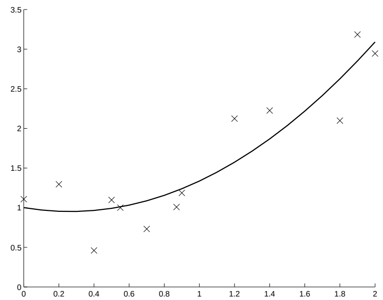
Example We want to fit a polynomial $p(t) = c_0 + c_1 t + c_2 t^2$ to a set or points (t_i, y_i), $1 \le i \le m$.



Formulate the over determined linear system and solve using the least squares method.

Suppose the data (t_i, x_i) is stored as two vectors t and x. Then

```
>> A=[ t.^0 t.^1 t.^2]; b=y; c=(A'*A)\(A'*b);
>> tt=0:0.1:2;yy=c(1)+c(2)*tt+c(3)*tt.^2;
>> plot(t,y,'x',tt,yy);
```



The QR decomposition

Definition A matrix Q is *orthogonal* if $Q^T Q = I$, where I is the identity.

Theorem Let $A \in \mathbb{R}^{m\times n}$, $m > n$. Then we can write

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

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

Properties R is non-singular if A has linearly independent columns and $Q = (q_1, q_2, \ldots, q_m)$ provides an orthonormal basis for $\mathbb{R}^m$.

How to use for solving least squares problems?

Lemma Suppose that Q is orthogonal and x is a vector. Then $\|Qx\|_2 = \|x\|_2$.

Theorem Let $A = Q_1 R$, $Q = (Q_1, Q_2)$. Then the x that minimize $\|Ax - b\|_2$ is given by

$$\|Ax - b\|_2^2 = \left\| \begin{pmatrix} R \\ 0 \end{pmatrix} x - \begin{pmatrix} Q_1^T b \\ Q_2^T b \end{pmatrix} \right\|_2^2,$$

and thus $x = R^{-1} Q_1^T b$ and $\|r\|_2 = \|Q_2^T b\|_2$.

We only need the *reduced QR* decomposition $A = Q_1 R$.

Lemma Computing R 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.

Algorithms for computing the QR decomposition is covered in *TANA15 Numerical Linear Algebra*.

In Matlab

```
>> [Q,R]=qr(A,0);
>> x=R\(Q'*b);
```

Lemma Suppose $A = Q_1 R$ is the reduced QR decomposition. Then $\mathrm{range}(A) = \mathrm{span}(Q_1)$.

Remarks We have an orthonormal basis for $\mathrm{range}(A)$. The least squares problem is solved by computing the projection of b onto $\mathrm{range}(A)$.

We avoid the high condition number $\kappa_2(A^T A) = \kappa_2(R)^2$. Much more stable numerically.

Circle fitting

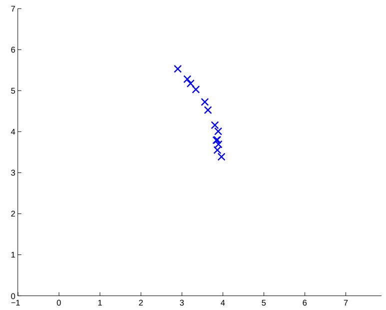
Example Points $(x, y)^T$ on a circle satisfies

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

where $u = (a, b_1, b_2)^T$ are the unknown parameters. The *center* and *radius* are given by

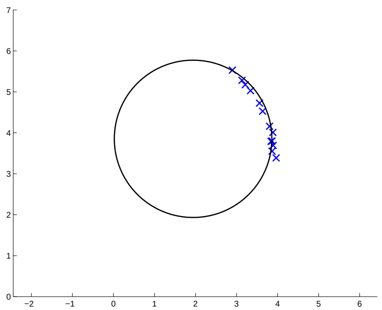
$$z = (x_0, y_0)^T = -(\frac{b_1}{2a}, \frac{b_2}{2a})^T, \quad \text{and} \quad r^2 = \frac{b_1^2 + b_2^2}{4a^2} - \frac{1}{a}.$$

Suppose we have two vectors x and y with $m = 13$ points on the circle. How to estimate the center and radius?



The points are stored in `x` and `y`. In Matlab we write

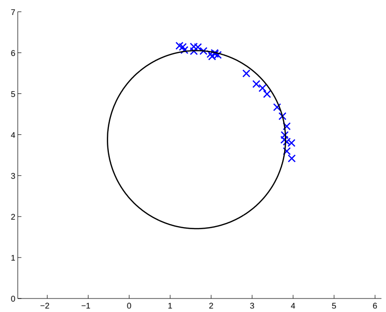
```
>> A=[x.^2+y.^2 x y]; b=-ones(size(x));
>> [Q,R]= qr( A ,0) );
>> u=R\(Q'*b);
>> z=-u(2:3)/2/u(1)
>> r=sqrt( (u(2)^2+u(3)^2)/4/u(1)^2-1/u(1))
```

The result is $z \approx (1.9400, 3.8533)^T$ and $r \approx 1.9195$. The exact circle is $z = (1.2, 3.4)^T$ and $r = 2.7$.

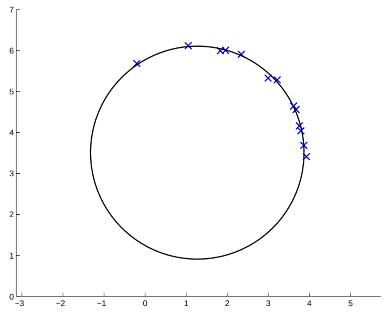
We need more data $(x_k, y_k)^T$ to improve the results.

The condition number is $\kappa_2(R) = 1274.4$.



Now we have $m = 25$ points that cover more of the circle. The condition number is $\kappa_2(R) = 722.7$.

The circle fit is $z \approx (1.6248, 3.8785)^T$ and $r \approx 2.1726$.



With $m = 13$ points more spread out across the circle we get $\kappa_2(R) = 403.5$.

The approximate circle is $z \approx (1.2707, 3.5083)^T$ and $r \approx 2.5980$.

If the points $(x_k, y_k)^T$ are close then all equations are roughly the same. A good set of observations is important!

Summary

- The condition number $\kappa_2(A)$ determines how sensitive $Ax = b$ is with respect to errors in the right hand side b or the matrix A.
- The least squares problem is to minimize $\|Ax - b\|_2$.
- Solve using the *normal equations* $A^T Ax = A^T b$. Mostly used for theoretical proofs.
- The decomposition $A = QR$ gives an orthogonal basis for $\text{range}(A)$. The standard method for solving least squares problems.
- Good standard software for computing the QR decomposition exists, e.g. in *ATLAS*.