Lecture 27 (hrs. 54,55) - November 26, 2025, 11:30 - 13:30 A13

(3.20) Examples.

(1) Find the least squares solutions of the system A x = b:

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

and the orthogonal projection of b onto C(A).

Solution: The system of the normal equations is: $3x = 1$ hence the unique least squares solution is: $x^* = 1/3$. The orthogonal projection of b onto C(A) is:

$$b_* = 1/3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

(2) Find the least squares solutions of the system $\underline{A}\, x = \underline{b}$:

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

and the orthogonal projection of b onto $C(\underline{A})$.

Solution: The system of the normal equations is:

$$\begin{bmatrix} 3 & 6 \\ 6 & 12 \end{bmatrix} x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

and:

$$\ker \underline{A}^t \underline{A} = \text{span}\{ \begin{bmatrix} 2 \\ -1 \end{bmatrix} \}$$

Once defined:

$$y = \begin{bmatrix} 1/3 \\ 0 \end{bmatrix}$$

we get:

$$S_{MQ}(\underline{A},\underline{b}) = \begin{bmatrix} 1/3 \\ 0 \end{bmatrix} + \text{span}\{ \begin{bmatrix} 2 \\ -1 \end{bmatrix} \} = \{ \begin{bmatrix} 1/3+2\lambda \\ -\lambda \end{bmatrix} , \lambda \in R \}$$

In this case, since the columns of $\underline{A}$ are *linearly dependent*, the set $S_{MQ}(\underline{A},\underline{b})$ has *infinite elements*.

The orthogonal projection of $\underline{b}$ onto $C(\underline{A})$ is:

$$\underline{b}_* = 1/3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Since $\underline{b}$ = b and C($\underline{A}$) = C(A), the orthogonal projection is the same as in the example. Above

(3.21) Example (weighted least squares).

Let A x = b be the system:

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

and $\underline{A}$ x = $\underline{b}$ be the system:

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

obtained by the system A x = b by taking two times the first and the third equations. The two systems are *equivalent* but: $S_{MQ}(A,b) \neq S_{MQ}(\underline{A},\underline{b})$. Indeed:

$$S_{MQ}(A,b) = \begin{bmatrix} 2/3 \\ 1/2 \end{bmatrix} \quad , \quad S_{MQ}(\underline{A},\underline{b}) = \begin{bmatrix} 5/9 \\ 1/2 \end{bmatrix}$$

This *should not be surprising*, in fact for the two systems we have $SQ(x) \neq \underline{SQ}(x)$:

$$SQ(x) = (x_1 - x_2)^2 + (x_1 - 1)^2 + (x_1 + x_2 - 1)^2$$

and:

$$\underline{SQ}(x) = 4\,(x_1 - x_2)^2 + (x_1 - 1)^2 + 4\,(x_1 + x_2 - 1)^2$$

The $\underline{SQ}$ function is obtained by *weighting* the addends of the SQ function with positive ‘weights’.

(3.22) Example.

Let: (-1,0), (0,1), (1,1) be given data. Find the *best-fitting elements* of F = span{ 1,x } to the data in the sense of least squares.

Observe that, given a Cartesian plane, each element of F has a non-vertical line as its graph. The problem can therefore be reformulated as: Find *the lines* that best fit the data in the least squares sense.

As shown in Example (3.14) and Remark (3.15) of Lecture 25, the problem is solved by determining the solutions in the least squares sense of the system (which translates the interpolation conditions):

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

The system of the normal equations has only one solution (indeed the columns...):

$$x^* = \begin{bmatrix} 2/3 \\ 1/2 \end{bmatrix}$$

and the only line that best fits the data in the least squares sense is the graph of the only element of F = span{ 1,x } that best fits the data in the least squares sense:

$$p(x) = \frac{2}{3} + \frac{1}{2}\,x$$

(3.23) Example.

Let (1,0), (1/2,1), (1/3,2) be given data. Find the elements of F = span{ 1,1/x } that best fit the data in the least squares sense.

Proceeding as in the previous example, the problem is solved by determining the solutions in the least squares sense of the system (which translates the interpolation conditions):

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} x = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

The system of the normal equations has only one solution :

$$x^* = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

and the only element of F that best fit the data in the least squares sense is:

$$f(x) = -1 + \frac{1}{x}$$

(3.24) Remark.

Let $A \in R^{r \times c}$ where r > c, let $b \in R^r$ and let $S_{MQ}(A,b)$ the set of the least squares solutions of A x = b. We have:[1]

there exists only one column $y_* \in S_{MQ}(A,b)$ of *least norm*[2]

(3.25) Example.

Let A x = b be the system:

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

It is:

1 Proof omitted.

2 Formally: there exists a unique point of absolute minimum of the function $\| x \|$ on $S_{MQ}(A,b)$.

$$S_{MQ}(A,b) = \begin{bmatrix} 1/2 \\ 0 \end{bmatrix} + \text{span}\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix} \}$$

and, by drawing the set $S_{MQ}(A,b)$ on a Cartesian plane, it is easy to verify that the element of minimum norm (i.e. the one closest to the origin) is:

$$y_* = \begin{bmatrix} 1/4 \\ 1/4 \end{bmatrix}$$

(3.26) Remark (pseudoinverse matrix).

Let $A \in R^{r \times c}$ where r > c. For each $b \in R^r$, let $S_{MQ}(A,b)$ be the set of the least squares solutions of A x = b.

The function $F: R^r \to R^c$ defined by:

$$F(b) = \text{the least norm element of } S_{MQ}(A,b)$$

is a *linear map* from R^r into R^c.[3] Hence there exists a $c \times r$ matrix, denoted by A^+, such that:

$$F(b) = A^+ b$$

The matrix A^+ is called *pseudoinverse matrix of* A.

If the columns of A are linearly independent then (see Remark (3.19) of Lecture 26) $S_{MQ}(A,b)$ has only one element, which is the one with minimum norm, and from the normal equations we obtain, since $A^t A$ is invertible:

$$F(b) = (A^t A)^{-1} A^t b$$

In this case we have then:

$$A^+ = (A^t A)^{-1} A^t$$

Note that if $A \in R^{n \times n}$ is an invertible matrix, then $A^+ = A^{-1}$. This explains why A^+ is called pseudoinverse matrix.

(3.27) Example.

Find the pseudoinverse matrix of:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$$

By definition, $A^+ \in R^{2 \times 3}$ is the only matrix such that: for every $b \in R^r$, $A^+ b$ = the least norm column in $S_{MQ}(A,b)$. Let e_1, e_2, e_3 be the columns of the canonical basis of R^3. Then the three columns of A^+ are:

$$F(e_1) \; , \; F(e_2) \; , \; F(e_3)$$

It is:

$$S_{MQ}(A,e_1) = \begin{bmatrix} 1/3 \\ 0 \end{bmatrix} + \text{span}\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix} \}$$

3 Proof omitted.

hence, proceeding as in Example (3.25):

$$F(e_1) = \begin{bmatrix} 1/6 \\ 1/6 \end{bmatrix}$$

Analogously we find:

$$F(e_2) = F(e_3) = \begin{bmatrix} 1/6 \\ 1/6 \end{bmatrix}$$

Finally:

$$A^+ = \begin{bmatrix} 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 \end{bmatrix}$$

(3.28) Definition (QR factorization: non-square matrices).

Let $A \in R^{r \times c}$ where $r > c$. A *QR factorization* of A is a pair U,T such that:

- $U \in R^{r \times c}$ is a matrix with orthonormal columns
- $T \in R^{c \times c}$ is an upper triangular matrix
- $U\,T = A$

(3.29) Example (use of the GS procedure to find a QR factorization of a non-square matrix).

Let:

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \in R^{3 \times 2}$$

To search for a QR factorization of A we may use a straightforward generalization of the GS procedure. Let a_1 and a_2 be the columns of A:

Step one.

We look for $\Omega = [\omega_1, \omega_2] \in R^{3 \times 2}$ with orthogonal columns and $\Theta \in R^{2 \times 2}$ upper triangular with $\theta_{kk} = 1$ such that $\Omega\,\Theta = A$. *If such matrices exist*, rewriting the last equality *column by column* we have:

$$\omega_1 = a_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad , \quad \omega_1\,\theta_{1,2} + \omega_2 = a_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

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

$$(\omega_1\,\theta_{1,2}) \bullet \omega_1 + \omega_2 \bullet \omega_1 = a_2 \bullet \omega_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \bullet \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 2$$

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

$$\theta_{1,2} = (a_2 \bullet \omega_1) \,/\, (\omega_1 \bullet \omega_1) = 2/3$$

hence:

$$\omega_2 = a_2 - \omega_1\,\theta_{1,2} = \begin{bmatrix} -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

Finally:

$$\Omega = \begin{bmatrix} 1 & -2/3 \\ 1 & 1/3 \\ 1 & 1/3 \end{bmatrix} \quad \text{and} \quad \Theta = \begin{bmatrix} 1 & 2/3 \\ 0 & 1 \end{bmatrix}$$

Step two.

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: $\Delta = \mathrm{diag}(\|\omega_1\|, \|\omega_2\|) = \mathrm{diag}(\mathrm{sqrt}(3), \mathrm{sqrt}(2/3))$. We easily check that the pair:

$$U = \Omega\,\Delta^{-1} = \begin{bmatrix} 1/\sqrt{3} & -\sqrt{2/3} \\ 1/\sqrt{3} & 1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{6} \end{bmatrix} \quad , \quad T = \Delta\,\Theta = \begin{bmatrix} \sqrt{3} & 2/\sqrt{3} \\ 0 & \sqrt{2/3} \end{bmatrix}$$

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

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

(3.30) Remark (QR factorization and least squares).

Let U,T be a QR factorization of $A \in R^{r \times c}$ where $r > c$. It is:

$$A^t A = (U\,T)^t (U\,T) = (T^t U^t)(U\,T) = T^t (U^t U)\,T$$

Since the matrix $U \in R^{r \times c}$ has orthonormal columns, then we have $U^t U = I \in R^{c \times c}$. Then:

$$A^t A = T^t T$$

Moreover:

$$A^t b = (U\,T)^t b = T^t U^t b$$

If the matrix T is invertible, that is, if the columns of A are linearly independent, then T^t is also invertible and the systems

$$A^t A\,x = A^t b \quad \text{and} \quad T\,x = U^t b$$

are *equivalent*. The two systems, however, *do not have the same conditioning properties*. In fact, we have:[4]

$$c_2(A^t A) = (c_2(T))^2$$

that is: the system $A^t A\,x = A^t b$ has *worse* conditioning properties than the system $T\,x = U^t b$. To determine the least squares solutions of $A\,x = b$ using a computer, we first find a QR factorization of A and then we solve the system $T\,x = U^t b$.

4 Proof omitted.