Lecture 15

We will apply the theory of least square solutions studied in the previous lecture to some specific problems.

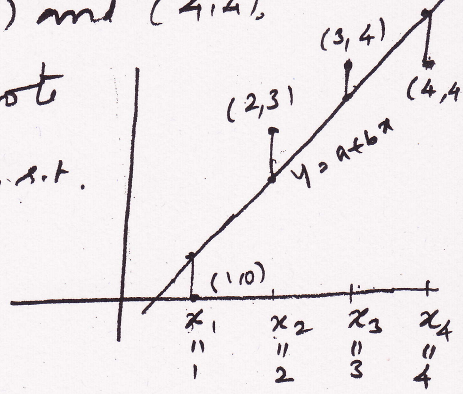
Least squares fitting: Let us start with the following example.

Example: Let us find a straight line $y = a + bx$ that fits best the given data $(1,0), (2,3), (3,4)$ and $(4,4)$.

The best is in the following sense: Let us denote the data by (x_i, y_i), $i = 1, 2, 3, 4$. Find $a, b \in \mathbb{R}$ s.t.

$$\sum_{i=1}^{4} |a + b x_i - y_i|^2 \quad \text{----} (*)$$

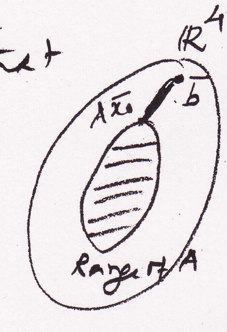
is minimum (see the figure).



We will now convert the minimization problem into a problem involving a matrix.(*) Let

$$A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}, \quad \bar{x} = \begin{bmatrix} a \\ b \end{bmatrix} \quad \& \quad \bar{b} = \begin{bmatrix} 0 \\ 3 \\ 4 \\ 4 \end{bmatrix}.$$

Here $A : \mathbb{R}^2 \to \mathbb{R}^4$ and $\bar{b} \in \mathbb{R}^4$. One can verify that $b \notin$ Range of A. The above problem $(*)$ can be written as: Find $\bar{x}_0 = (a_0, b_0) \in \mathbb{R}^2$ such that



$$\| A\bar{x}_0 - \bar{b} \|^2 = \min_{\bar{x} \in \mathbb{R}^2} \| A\bar{x} - \bar{b} \|^2$$

So $\bar{x}_0$ is a least square solution & it is a solution of the normal equation $A^T A \bar{x}_0 = A^T \bar{b}$, ie $\bar{x}_0 = (A^T A)^{-1}(A^T \bar{b})$ if the inverse exists. Here the inverse exists &

$$(A^T A) = \begin{bmatrix} 4 & 10 \\ 10 & 30 \end{bmatrix}, \quad (A^T A)^{-1} = \begin{bmatrix} 3/2 & -1/2 \\ -1/2 & 1/5 \end{bmatrix}, \quad A^T \bar{b} = \begin{bmatrix} 11 \\ 34 \end{bmatrix}$$

$$\& \quad \bar{x}_0 = (A^T A)^{-1} A^T \bar{b} = \begin{bmatrix} -1/2 \\ 13/10 \end{bmatrix}.$$

Therefore $y = -\frac{1}{2} + \frac{13}{10} x$ is the line that we are looking for.

(*) First observe that for $f(x) = a + bx$, the data lead to an inconsistant system $A\bar{x} = \bar{b}$.

Remark: In the previous example, we have used a linear function to approximate the given data. One can also use the same method to approximate by a quadratic polynomial or a polynomial of degree any $k \geq 1$.

Diagonalization:

If a system $Ax = b$ is consistent, we try to solve it by Gauss elimination method. If it is not consistent, then we find a least square solution by solving the normal equation $A^TAx = A^Tb$. If A^TA is invertible then one can write the solution explicitly: $x = (A^TA)^{-1}A^Tb$. Of course, here one can ask the question: under what assumption on A, the matrix A^TA is invertible. Note that A^TA is a symmetric square matrix. One can also use this additional property to check whether the matrix A^TA is invertible or not. For example at the end of this course we prove the following result:

If A is a (real) symmetric matrix, then there exists an invertible matrix Q s.t. $A = QDQ^{-1}$ where D is a diagonal matrix (or $Q^{-1}AQ = D$).

In this case, we say that A is diagonalizable. To show that A is invertible, we have to show that D is invertible which is easy. Note that evaluating the determinant is also relatively easier. Moreover, if A is diagonalizable, evaluating A^k is also easy. For example, $A^2 = (QDQ^{-1})(QDQ^{-1}) = QD^2Q^{-1}$ & $A^k = QD^kQ^{-1}$. In many applications we deal with A^k, for large k; esspecially, when we use some recursive method. We will see an example.

In order to study the diagonalization, we introduce two concepts eigenvalues and eigenvectors, which play important roles in their own right in Mathematics and have applications in other fields.

Definition: Let A be an $n\times n$ matrix. A nonzero vector $x\in\mathbb{R}^n$ is called an eigenvector (or characteristic vector) of A if $\exists$ $\lambda\in\mathbb{R}$ s.t. $Ax=\lambda x$. The scalar λ is called an eigenvalue (or characteristic value) of A.

Geometrically: x & Ax are parallel.

Algebraically: $x\neq 0$ s.t. $(\lambda I-A)x=0$ ie $x\in N(\lambda I-A)$.

How to find x & λ: First note that

λ is an eigenvalue of A $\Leftrightarrow$ $(\lambda I-A)x=0$ has a nontrivial sol

$\Longleftrightarrow$ $\det(\lambda I-A)=0$ which is a polynomial of degree n in λ.

Therefore, the eigenvalues are just the roots of the characteristic equation $\det(\lambda I-A)=0$ or $\det(A-\lambda I)=0$

For each eigenvalue the space $N(\lambda I-A)$, the collection of eigenvectors, is called the eigenspace corresponding to λ.

Example: (Matrix with distinct eigenvalues): Let $A=\begin{bmatrix}4&-5\\2&-3\end{bmatrix}$

Then $|A-\lambda I|=0 \Rightarrow \begin{vmatrix}4-\lambda & -5\\ 2 & -3-\lambda\end{vmatrix}=0 \Rightarrow \lambda^2-\lambda-2=0 \Rightarrow \lambda_1=-1$ & $\lambda_2=2$. are the e-values

Eigen space for $\lambda_1=-1$:

$(A-\lambda_1 I)X=0 \Rightarrow \begin{bmatrix}5&-5\\2&-2\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix} \Rightarrow x_1=x_2 \Rightarrow X_1=(1,1)$ is an eigenvector &

the eigen space corresponding to $\lambda_1=-1$ is $E(\lambda_1)=\{t(1,1): t\in\mathbb{R}\}$.

Eigen space for $\lambda_2=2$:

$(A-\lambda_2 I)X=0 \Rightarrow \begin{bmatrix}2&-5\\2&-5\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix} \Rightarrow 2x_1=5x_2 \Rightarrow X_2=(5,2)$ is an eigen vector &

The eigen space corresponding to $\lambda_2=2$ is $E(\lambda_2)=\{tX_2: t\in\mathbb{R}\}$.

Note that $\{X_1, X_2\}$ is L.I.