lecture 15

We will apply the theory of least square solutions studied in its 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 \( a, b \in \mathbb{R} \) such that

\[
\sum_{i=1}^{4} |a + bx_i - y_i|^2 \rightarrow \min
\]

is minimum. (See figure 1.)

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

\[
A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_4 \end{bmatrix}, \quad \bar{x} = \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{and} \quad \bar{b} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_4 \end{bmatrix}
\]

Here \( A : \mathbb{R}^2 \rightarrow \mathbb{R}^4 \) and \( \bar{b} \in \mathbb{R}^4 \). One can verify that \( \bar{b} \) is in the 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 and it is a solution of the normal equation \( A^T A \bar{x}_0 = A^T \bar{b} \), i.e.,

\[
\bar{x}_0 = (A^T A)^{-1} A^T \bar{b}
\]

If the inverse exists, then the inverse exists and

\[
(\bar{x}_0) = \begin{bmatrix} 4 & 10 \\ 10 & 30 \end{bmatrix}, \quad (\bar{x}_0)^{-1} = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{5} \end{bmatrix}, \quad A^T \bar{b} = \begin{bmatrix} 3/4 \\ 1/10 \end{bmatrix},
\]

\[
\bar{x}_0 = (A^T A)^{-1} A^T \bar{b} = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}
\]

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

(\star) First observe that for \( f(x) = a + bx \), the data lead to an inconsistent system \( A \bar{x} = \bar{b} \).
Remark: In the previous example, we have used a linear function to approximate 10 given data. One can also use the same method to approximate by a quadratic polynomial or a polynomial of degree any k > 1.

Diagonalization:

If a system $A x = b$ is consistent, we try to solve it by Gaussian elimination method. If it is not consistent, then we find a least square solution by solving the normal system $A^T A x = A^T b$. If $A^T A$ is invertible then one can write the solution explicitly: $x = (A^T A)^{-1} A^T b$. Of course, here one can ask the question: under what assumption on $A$, the matrix $A^T A$ is invertible. Note that $A^T A$ is a symmetric square matrix. One can also use this additional property to check whether the matrix $A^T A$ 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. $A = Q D Q^{-1}$ where $D$ is a diagonal matrix (or $Q^{-1} A Q = D$).

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, $A^2 = (Q D Q^{-1})(Q D Q^{-1}) = Q D^2 Q^{-1}$ and $A^k = Q D^k Q^{-1}$.

In many applications we deal with $A^k$, for large $k$, especially, 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 non-zero vector $x \in \mathbb{R}^n$ is called an eigenvector (or characteristic vector) of $A$ if there exists a scalar $\lambda \in \mathbb{R}$ such that $Ax = \lambda x$. The scalar $\lambda$ is called an eigenvalue (or characteristic value) of $A$.

**Geometrically:** $x$ and $Ax$ are parallel.

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

**How to find $x$ & $\lambda$:** First note that

$\lambda$ is an eigenvalue of $A$ $\iff$ $(\lambda I - A)x = 0$ has a non-trivial solution $x$.

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

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

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

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

**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.

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

**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 eigenvector.

**Eigen space corresponding to $\lambda_2 = 2$:** $E(\lambda_2) = \{ t(5,2): t \in \mathbb{R} \}$.

Note that $\{x_1, x_2\}$ is L.I.