Linear Algebra  Linear Least Squares  Lecture 27

Last Time: Orthogonal projections

This Time: Use projection to solve approximation problems

Given \( U \subseteq V \) a finite dimensional subspace, \( V = U \oplus U^\perp \)

\( \forall v \in V \) can be written as \( v = u + w \), \( u \in U \), \( w \in U^\perp \)

\( P_u(v) = \sum_{j=1}^{m} <v, f_j> f_j \)

\( P_u^\perp(v) = v - P_u(v) \)

If \( \{f_1, \ldots, f_m\} \) is an orthonormal basis of \( U \), \( P_u(v) = \sum_{j=1}^{m} <v, f_j> f_j \)

\[ \begin{bmatrix} P_u \end{bmatrix}_E = \begin{bmatrix} \sum_{j=1}^{m} f_j f_j^* \end{bmatrix} \]

For an arbitrary basis \( \{g_1, \ldots, g_m\} \) of \( U \), we set

\[ A = \begin{bmatrix} g_1 & \cdots & g_m \end{bmatrix} \] and \[ \begin{bmatrix} P_u \end{bmatrix}_E = A (A^T A)^{-1} A^T \]

Then \( U \subseteq V \) finite dimensional subspace

i) For every \( v \in V \), \( \|P_u(v)\| \leq \|v\| \) equality if and only if \( v \in U \)

ii) \( P_u(v) \) is the point in \( U \) closest to \( v \):

\[ \|v - P_u(v)\| \leq \|v - u\| \] \( \forall u \in U \) equality only if \( u = P_u(v) \)

Proof:

i) \[ \|v\|^2 = \|P_u(v) + P_u^\perp(v)\|^2 = \|P_u(v)\|^2 + \|P_u^\perp(v)\|^2 \geq \|P_u(v)\|^2 \]

with equality if and only if \( \|P_u^\perp(v)\|^2 = 0 \) \( \Leftrightarrow v = P_u(v) \) \( \forall v \in U \)

iii) For any \( u \in U \)

\[ \|v - u\|^2 = \|v - P_u(v) + P_u(v) - u\|^2 = \|P_u^\perp(v) + P_u(v) - u\|^2 = \|P_u^\perp(v)\|^2 + \|P_u(v) - u\|^2 \geq \|P_u^\perp(v)\|^2 = \|v - P_u(v)\|^2 \]

with equality if and only if \( \|P_u^\perp(v)\|^2 = 0 \) \( \Leftrightarrow v = P_u(v) \)
Linear Least Squares:

Suppose we have a data set \( \{(x_i, y_i)\}_{i=1}^{n} \subseteq \mathbb{R}^n \)
We want to find the line \( y = mx + b \)

which comes closest to matching these points.

Each choice of \( m \) \& \( b \) would give points \( \{(x_i, mx_i + b)\}_{i=1}^{n} \)
We want to choose \( m \) \& \( b \) to minimize the error \( \sum_{i=1}^{n} (mx_i + b - y_i)^2 \)

One way to think about this is to define
\[
U = \left\{ \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} : m, b \in \mathbb{R} \right\} = \left\{ mx + b : m, b \in \mathbb{R} \right\} \subseteq \mathbb{R}^n, \text{ where } x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}
\]

This is the two-dimensional subspace \( \text{span}(x, 1) \) of \( \mathbb{R}^n \)

The distance of the pt \( \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \) to a point \( mx + b \in U \)

is \( \sqrt{\sum_{i=1}^{n} (mx_i + b - y_i)^2} \).

So we've recast the original problem as:
find the point of \( U \) that is closest to \( y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \).
We know the solution to this is \( P_U(y) \).

E.g., Find the line in the plane that comes closest to containing
the data points \((-2, 1), (-1, 2), (0, 5), (1, 4), (2, 8)\).

Here \( x = \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \end{bmatrix}, \ 1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \ U = \mathbb{R}^2 \cap \text{span}(x, 1), \ y = \begin{bmatrix} 1 \\ 5 \\ 8 \end{bmatrix} \). \( P_U(y) \) ?

Let \( A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \) so \( [P_U]_E = A (A^T A)^{-1} A^T = \frac{1}{10} \begin{bmatrix} 1 & 4 & -1 & 0 & 2 \\ 4 & 16 & 0 & 2 & 8 \\ -1 & 0 & 1 & 4 & 2 \\ 0 & 2 & 4 & 1 & 16 \\ 2 & 8 & 2 & 16 & 4 \end{bmatrix} \)

\( A [P_U]_E [y]_E = \frac{1}{10} \begin{bmatrix} 8 \\ 26 \\ 56 \\ 88 \end{bmatrix} = A [y]_E \)

So the line \( y = \frac{8}{3} x + 4 \) is the best fit to the data.
E.g., find the closest quadratic polynomial to $e^x$ on $[0,1]$ with distance induced by the inner product $\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx$.

We want to find $a, b, c \in \mathbb{R}$ making $\int_0^1 (e^x - (a + bx + cx^2))^2\,dx$ as small as possible.

As we've seen, the basis $(1, x, x^2)$ of $P_2(\mathbb{R})$ is not orthonormal for this inner product. Using Gram-Schmidt, we find the orthonormal basis $(1, 2\sqrt{3}x - \sqrt{3}, 6\sqrt{3}x^2 - 6\sqrt{3}x + \sqrt{3}) = (f_1, f_2, f_3)$

We compute
\[
\begin{align*}
\langle e^x, f_1 \rangle &= e - 1 \\
\langle e^x, f_2 \rangle &= -\sqrt{3}e + 3\sqrt{3} \\
\langle e^x, f_3 \rangle &= 7\sqrt{3}e - 19\sqrt{3} \\
\end{align*}
\]

so $P_{e(x)}(e^x) = (e-1)f_1 + (-\sqrt{3}e + 3\sqrt{3})f_2 + (7\sqrt{3}e - 19\sqrt{3})f_3$

Another way of thinking about this is: Given $A \in M_{m \times n}(\mathbb{F})$, if $b \in \mathbb{F}^n$ is not in $\text{C}((A))$, then $Ax = b$ doesn't have any solution.

We can solve $Ax = P_{\text{C}(A)}(b)$.

Indeed, $Ax = P_{\text{C}(A)}(b) = A(A^*A)^{-1}A^*b$ if columns of $A$ linearly independent.

so the solution is $x = (A^*A)^{-1}A^*b$.

The best approximate solution to $Ax = b$ is $x = (A^*A)^{-1}A^*b$ (assuming columns of $A$ linearly independent).
E.g., Find the best-fit ellipse to the points
(0,2), (2,1), (1,-1), (-1,2), (-3,1), (-1,-1)

Here the general equation for an ellipse (or any conic section) is
\[ x^2 + B y^2 + C x y + D x + E y + F = 0 \]
If the points were on an ellipse we would have
\[ 0^2 + B (4) + C (0)(2) + D (0) + E (2) + F = 0 \]
\[ 2^2 + B (1)^2 + C (1)(1) + D (2) + E (1) + F = 0 \]
\[ \vdots \]
\[ (1)^2 + B (-1)^2 + C (-1)(1) + D (-1) + E (-1) + F = 0 \]

\[ \begin{bmatrix} 4 & 0 & 0 & 0 & 0 & 0 \\
1 & -1 & 0 & 0 & 0 & 0 \\
0 & 1 & -1 & 0 & 0 & 0 \\
0 & 0 & 1 & -1 & 0 & 0 \\
0 & 0 & 0 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} B \\
C \\
D \\
E \\
F \end{bmatrix} = \begin{bmatrix} 0 \\
0 \\
0 \\
0 \\
0 \\
0 \end{bmatrix} \]

\[ \begin{bmatrix} 405/1336 \\
-89/133 \\
201/133 \\
-123/133 \\
-617/133 \end{bmatrix} = (A^T A)^{-1} A^T b \]

So we look for the best approximate solution to \( Ax = b \)

This would be \( x = (A^T A)^{-1} A^T b \)

\[ x = \begin{bmatrix} 405/1336 \\
-89/133 \\
201/133 \\
-123/133 \\
-617/133 \end{bmatrix} \]

i.e., the best-fit ellipse is
\[ x^2 + \frac{405}{1336} y^2 - \frac{89}{133} x y + \frac{201}{133} x - \frac{123}{133} y - \frac{617}{133} = 0 \]