Lecture #25 LEAST SQUARES

Consider the linear system

$$L(u) = v$$

where $L : U \xrightarrow{\text{linear}} W$

and $v \in W$ is given.

So u are the <u>unknowns</u>.

Sometimes such a system has a unique solutions, sometimes there are many solutions (a solution space). But if v is <u>not</u> in the range of L

$$v \notin L(W)$$

we will <u>never</u> solve $L(u) = v$ for u. i.e. this system is consistent if and only if $v \in L(W)$.

But for many problems, we do not need an <u>exact</u> solution and nor do we care if an exact solution even exists. Instead we simply want to find the <u>best</u> approximation possible, i.e. find u such that

$$L(u) - v$$

is as "<u>small as possible</u>".

"My work always tried to unite the Truth with the Beautiful, but when I had to choose one or the other, I usually chose the Beautiful"

Hermann Weyl

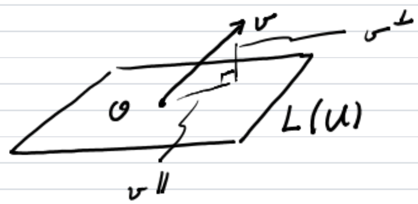
Suppose M is the matrix of L in bases for U & W while v and u are given by column vectors V and X in those bases.

Therefore we need to approximate

$$MX - V \approx 0 \quad \text{—} (*)$$

Note, if $\dim U = n$, $\dim W = m$ then M is an $m \times n$ *matrix*.

Idea



Write $W = L(U) \oplus L(U)^{\perp}$

$\Rightarrow \quad v = v^{\parallel} + v^{\perp}$ uniquely.

solve $L(u) = v^{\parallel}$

In components, $v^{\perp}$ is just $V - MX$

In terms of M *notice* that $L(V)$ is spanned by the rows of M (choose the natural basis $e_1, \dots, e_n$ for V, then $\{Me_1, \dots, Me_n\}$ are the rows of M). So $v^{\perp}$ must be perpendicular to the rows of M. i.e., $M^T(V - MX) = 0$ or

$$\boxed{M^T M X = M^T V}$$

Solutions X to $M^T M X = M^T V$ are called *least squares* solutions to $MX = V$.

Note any solution to $MX = V$ is a least squares solution but the converse is often false. (In fact $MX = V$ may have *no* solutions at all but still have *least squares* solutions to $M^T M X = M^T V$)

Suppose $\ker L = \{0\}$

ie the only solution to

$$Mx = 0$$

is the zero vector.

Then if $x \neq 0$ $\quad Mx \neq 0$.

But $M^T M x \neq 0$ also because if it were then

$$0 = x^T M\, Mx = |Mx|^2$$

or equivalently $Mx = 0$, a contradiction. Thus if $\ker L = \{0\}$, the matrix

$$M^T M$$

is <u>invertible</u> (because 0 is the only solution to $Mx = 0$). Therefore the least squares solution is unique and equals

$$\boxed{X = (M^T M)^{-1} M^T V}$$

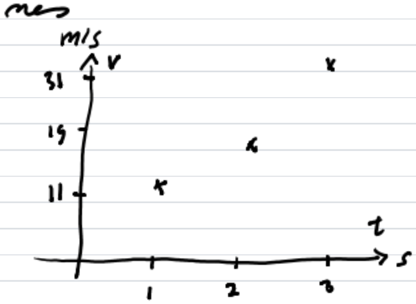
This is the least squares method.

① Compute $M^T M$ and $M^T V$

② Solve $(M^T M) X = (M^T V)$ by Gaussian elimination.

<u>Example</u> Captain Conundrum

falls off the leaning tower of Pisa and makes 3 (rather shaky) measurements of his velocity at 3 different times



He believes that his data are best approximated by a straight line

$$v = at + b$$

Which a and b best fit his data:

$$\begin{aligned} 11 &= a\cdot 1 + b \\ 19 &= a\cdot 2 + b \\ 31 &= a\cdot 3 + b \end{aligned} \iff \underbrace{\begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix}}_{M} \underbrace{\begin{pmatrix} a \\ b \end{pmatrix}}_{X} \overset{?}{=} \underbrace{\begin{pmatrix} 11 \\ 19 \\ 31 \end{pmatrix}}_{V}$$

There is likely NO straight line solution so instead solve $M^T M X = M V$

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 11 \\ 19 \\ 31 \end{pmatrix}$$

$$\Rightarrow \left(\begin{array}{cc|c} 14 & 6 & 142 \\ 6 & 3 & 61 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & 10 \\ 0 & 1 & 1/3 \end{array}\right)$$

The least squares fit is

$$v = 10t + \tfrac{1}{3}$$

Review Exercises

1. Let $L : U \to V$ be a linear transformation. Suppose $v \in L(U)$ and you have found a vector u_{ps} that obeys $L(u_{ps}) = v$

   Explain why you need to compute

   $$\ker L$$

   to find the solution space of the linear system

   $$L(u) = v.$$

④ Explain why elements of span $U \cup V$ all have the form $u + v$ with $u \in U$, $v \in V$.

⑤ Show that elements of $U \oplus V$ can be expressed uniquely as $u + v$ with $u \in U$, $v \in V$.
[Hint: suppose the opposite were true and derive a contradiction]

⑥ Use the subspace theorem to show that $U^\perp$ is a subspace.

(viii) Explain why $\lambda = \bar{\lambda}$. What does this mean?