Lecture 12: Least squares, OD II

Amos Ron

University of Wisconsin - Madison

February 26, 2021

Outline

1. Overdetermined system
   - The characterizaton theorem
   - The Normal Equation
   - The normal equation algorithm is unstable

Outline

1. Overdetermined system
   - The characterizaton theorem
   - The Normal Equation
   - The normal equation algorithm is unstable

Blank page

The abstract problem

Definition: Least square approximation in vector spaces

- V is a vector space (for example, R^m).
- W is a subspace of V.
- v is some vector in V.
- Find: $w^* \in W$ such that

$$||v - w^*||_2 \leq ||v - w||_2, \quad \forall w \in W.$$

The abstract problem

The characterization theorem

V, W, v as before. Assume $\widetilde{w} \in W$ satifying

$$v - \widetilde{w} \perp W.$$

Then $\widetilde{w}$ is the only solution to the least squares problem.

The abstract problem

The characterization theorem

V, W, v as before. Assume $\widetilde{w} \in W$ satifying

$$v - \widetilde{w} \perp W.$$

Then $\widetilde{w}$ is the only solution to the least squares problem.

Proof: Let $w \in W$, different from $\widetilde{w}$. We need to show that

$$||v - \widetilde{w}||_2^2 < ||v - w||_2^2.$$

We write:

$$||v-w||_2^2 = ||(v-\widetilde{w})+(\widetilde{w}-w)||_2^2 = ||v-\widetilde{w}||_2^2+||\widetilde{w}-w||_2^2 > ||v-\widetilde{w}||_2^2.$$

The middle equality since $\widetilde{w} - w \in W$, hence $(v - \widetilde{w}) \perp (\widetilde{w} - w)$, by assumption. The inquality $>$ is since $||\widetilde{w} - w||_2^2$ is positive, since we assume w is different from $\widetilde{w}$.

Back to the matrix formulation

How do we practically solve such an abstract problem?
Assume $V = R^m$. Usually, W is given in terms of $n < m$ vectors

$$w_1, \ldots, w_n$$

that form a basis for W.
Let $A_{m \times n}$ be the concatenation of the W-basis. Then W is the range of A.
Instead of looking for $w^* \in W$, such that $||w^* - v||_2$ is minimal, we look for $x^* \in R^n$ such that

$$||Ax^* - v||_2$$

is minimal:

$$w^* = \sum_{i=1}^{n} x^*(i) w_i.$$

Back to the matrix formulation

Suppose that the matrix problem is the original: $A_{m\times n}, b$ are given and we look for $x^* \in R^n$ such that

$$||Ax^* - b||_2$$

is minimal. Then

$$W = \mathrm{range}(A),$$

hence the columns of A span W (Normally, they form a basis for W).

Back to the matrix formulation

Suppose that the matrix problem is the original: $A_{m\times n}, b$ are given and we look for $x^* \in R^n$ such that

$$||Ax^* - b||_2$$

is minimal. Then

$$W = \text{range}(A),$$

hence the columns of A span W (Normally, they form a basis for W).
So, we look for $x^* \in R^n$, such that

$$(Ax^* - b, Ax) = 0, \quad \forall x \in R^n.$$

Then:

$$0 = (Ax^* - b, Ax) = (A'(Ax^* - b), x), \quad x \in R^n.$$

Back to the matrix formulation

So, we look for $x^* \in R^n$, such that

$$(Ax^* - b, Ax) = 0, \quad \forall x \in R^n.$$

Then:

$$0 = (Ax^* - b, Ax) = (A'(Ax^* - b), x), \quad x \in R^n.$$

The Normal Equation solution to the matrix version of the OD problem

- Every solution of the normal equation

$$A'Ax^* = A'b.$$

- The normal equation always have solutions, even in case $A'A$ is singular.

How do we know that the normal equation has solutions?

Since $\mathrm{range} A'A \subset \mathrm{range}(A')$, we just need to prove that

$$\mathrm{rank}(A'A) = \mathrm{rank}(A').$$

Since

$$\mathrm{rank}(A') = \mathrm{rank}(A),$$

we may prove

$$\mathrm{rank}(A'A) = \mathrm{rank}(A).$$

Let $W := \mathrm{range}(A)$. $\dim W = \mathrm{rank}(A)$. Let B be the restriction of A' to W. Then:

$$\mathrm{rank}(A'A) = \mathrm{rank}(B)$$

.
But,

$$\mathrm{rank}(B) = \dim W - \dim \ker(B) = \mathrm{rank}(A) - \dim \ker B.$$

How do we know that the normal equation has solutions?

But,

$$\mathrm{rank}(B) = \dim W - \dim\ker(B) = \mathrm{rank}(A) - \dim\ker B.$$

So, we need $\dim\ker B = 0$, i.e., $\ker B = \{0\}$: Suppose $A'w = 0$ for some $w \in W$. Then $A'Ax = 0$, for some $x \in R^n$. Need to show $Ax = 0$:

How do we know that the normal equation has solutions?

But,

$$\mathrm{rank}(B) = \dim W - \dim\ker(B) = \mathrm{rank}(A) - \dim\ker B.$$

So, we need $\dim\ker B = 0$, i.e., $\ker B = \{0\}$: Suppose $A'w = 0$ for some $w \in W$. Then $A'Ax = 0$, for some $x \in R^n$. Need to show $Ax = 0$:

$$0 = (A'Ax, x) = (Ax, Ax) = ||Ax||_2^2 \implies Ax = 0.$$

The instability issue

In the 2-norm

$$\mathrm{cond}(A'A) = \mathrm{cond}(A)^2.$$

What to do?