2019-10-09

1 Trouble points

At a high level, there are two pieces to solving a least squares problem:

1. Project b onto the span of A.
2. Solve a linear system so that Ax equals the projected b.

Consequently, there are two ways we can get into trouble in solving least squares problems: either b may be nearly orthogonal to the span of A, or the linear system might be ill conditioned.

1.1 Perpendicular problems

Let's first consider the issue of b nearly orthogonal to the range of A first. Suppose we have the trivial problem

$$A = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad b = \begin{bmatrix} \epsilon \\ 1 \end{bmatrix}.$$

The solution to this problem is $x = \epsilon$; but the solution for

$$A = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \hat{b} = \begin{bmatrix} -\epsilon \\ 1 \end{bmatrix}.$$

is $\hat{x} = -\epsilon$. Note that $\|\hat{b} - b\|/\|b\| \approx 2\epsilon$ is small, but $|\hat{x} - x|/|x| = 2$ is huge. That is because the projection of b onto the span of A (i.e. the first component of b) is much smaller than b itself; so an error in b that is small relative to the overall size may not be small relative to the size of the projection onto the columns of A.

Of course, the case when b is nearly orthogonal to A often corresponds to a rather silly regressions, like trying to fit a straight line to data distributed uniformly around a circle, or trying to find a meaningful signal when the signal to noise ratio is tiny. This is something to be aware of and to watch out for, but it isn't exactly subtle: if $\|r\|/\|b\|$ is near one, we have a numerical problem, but we also probably don't have a very good model.

1.2 Conditioning of least squares

A more subtle problem occurs when some columns of A are nearly linearly dependent (i.e. A is ill-conditioned). The *condition number of A for least squares* is

$$\kappa(A) = \|A\|\|A^\dagger\| = \sigma_1/\sigma_n.$$

If $\kappa(A)$ is large, that means:

1. Small relative changes to A can cause large changes to the span of A (i.e. there are some vectors in the span of $\hat{A}$ that form a large angle with all the vectors in the span of A).

2. The linear system to find x in terms of the projection onto A will be ill-conditioned.

If θ is the angle between b and the range of A, then the sensitivity to perturbations in b is

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\kappa(A)}{\cos(\theta)} \frac{\|\delta b\|}{\|b\|}$$

while the sensitivity to perturbations in A is

$$\frac{\|\delta x\|}{\|x\|} \leq \left(\kappa(A)^2 \tan(\theta) + \kappa(A)\right) \frac{\|\delta A\|}{\|A\|}.$$

The first term (involving $\kappa(A)^2$) is associated with the tendency of changes in A to change the span of A; the second term comes from solving the linear system restricted to the span of the original A. Even if the residual is moderate, the sensitivity of the least squares problem to perturbations in A (either due to roundoff or due to measurement error) can quickly be dominated by $\kappa(A)^2 \tan(\theta)$ if $\kappa(A)$ is at all large.

In regression problems, the columns of A correspond to explanatory factors. For example, we might try to use height, weight, and age to explain the probability of some disease. In this setting, ill-conditioning happens when the explanatory factors are correlated — for example, weight might be well predicted by height and age in our sample population. This happens reasonably often. When there is a lot of correlation, we have an *ill-posed* problem.

2 Sensitivity details

Having given a road-map of the main sensitivity result for least squares, we now go through some more details.

2.1 Preliminaries

Before continuing, it is worth highlighting a few facts about norms of matrices that appear in least squares problems.

1. In the ordinary two-norm, $\|A\| = \|A^T\|$.
2. If $Q \in \mathbb{R}^{m\times n}$ satisfies $Q^T Q = I$, then $\|Qz\| = \|z\|$. We know also that $\|Q^T z\| \leq \|z\|$, but equality will not hold in general.
3. Consequently, if $\Pi = QQ^T$, then $\|P\| \leq 1$. Equality actually holds unless Q is square (so that $\Pi = 0$).
4. If $A = QR = U\Sigma V^T$ are economy decompositions, then $\|A\| = \|R\| = \sigma_1(A)$ and $\|A^\dagger\| = \|R^{-1}\| = 1/\sigma_n(A)$.

2.2 Warm-up: $y = A^T b$

Before describing the sensitivity of least squares, we address the simpler problem of sensitivity of matrix-vector multiply. As when we dealt with square matrices, the first-order sensitivity formula looks like

$$\delta y = \delta A^T b + A^T \delta b$$

and taking norms gives us a first-order bound on absolute error

$$\|\delta y\| \leq \|\delta A\| \|b\| + \|A\| \|\delta b\|.$$

Now we divide by $\|y\| = \|A^T b\|$ to get relative errors

$$\frac{\|\delta y\|}{\|y\|} \leq \frac{\|A\| \|b\|}{\|A^T b\|} \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right).$$

If A were square, we could control the multiplier in this relative error expression by $\|A\| \|A^{-1}\|$. But in the rectangular case, A does not have an inverse. We can, however, use the SVD to write

$$\frac{\|A\| \|b\|}{\|A^T b\|} \geq \frac{\sigma_1(A) \|b\|}{\sigma_n(A) \|U^T b\|} = \kappa(A) \frac{\|b\|}{\|U^T b\|} = \kappa(A) \sec(\theta)$$

where $\theta \in [0, \pi/2]$ is the acute angle between b and the range space of A (or, equivalently, of U).

2.3 Sensitivity of the least squares solution

We now take variations of the normal equations $A^T r = 0$:

$$\delta A^T r + A^T(\delta b - \delta A x - A \delta x) = 0.$$

Rearranging terms slightly, we have

$$\delta x = (A^T A)^{-1} \delta A^T r + A^\dagger(\delta b - \delta A x).$$

Taking norms, we have

$$\|\delta x\| \leq \frac{\|\delta A\| \|r\|}{\sigma_n(A)^2} + \frac{\|\delta b\| + \|\delta A\| \|x\|}{\sigma_n(A)}.$$

We now note that because Ax is in the span of A,

$$\|x\| = \|A^\dagger A x\| \geq \|Ax\| / \sigma_1(A)$$

and so if θ is the angle between b and $\mathcal{R}(A)$,

$$\begin{aligned} \frac{\|b\|}{\|x\|} &\leq \sigma_1(A) \frac{\|b\|}{\|Ax\|} = \sigma_1(A) \sec(\theta) \\ \frac{\|r\|}{\|x\|} &\leq \sigma_1(A) \frac{\|r\|}{\|Ax\|} = \sigma_1(A) \tan(\theta). \end{aligned}$$

Therefore, we have

$$\frac{\|\delta x\|}{\|x\|} \leq \kappa(A)^2 \frac{\|\delta A\|}{\|A\|} \tan(\theta) + \kappa(A) \frac{\|\delta b\|}{\|b\|} \sec(\theta) + \kappa(A) \frac{\|\delta A\|}{\|A\|}.$$

which we regroup as

$$\frac{\|\delta x\|}{\|x\|} \leq \left(\kappa(A)^2 \tan(\theta) + \kappa(A) \right) \frac{\|\delta A\|}{\|A\|} + \kappa(A) \sec(\theta) \frac{\|\delta b\|}{\|b\|}.$$

2.4 Residuals and rotations

Sometimes we care not about the sensitivity of x, but of the residual r. It is left as an exercise to show that

$$\frac{\|\Delta r\|}{\|b\|} \leq \frac{\|\Delta b\|}{\|b\|} + \|\Delta \Pi\|$$

where we have used capital deltas to emphasize that this is not a first-order result: Δb is a (possibly large) perturbation to the right hand side and $\Delta \Pi = \hat{\Pi} - \Pi$ is the difference in the orthogonal projectors onto the spans of $\hat{A}$ and A. This is slightly awkward, though, as we would like to be able to relate the changes to the projector to changes to the matrix A. We can show[1] that $\|\Delta \Pi\| \leq \sqrt{2}\|E\|$ where $E = (I - QQ^T)\hat{Q}$. To finish the job, though, we will need the perturbation theory for the QR decomposition (though we will revert to first-order analysis in so doing).

Let $A = QR$ be an economy QR decomposition, and let $Q_\perp$ be an orthonormal basis for the orthogonal complement of the range of Q. Taking variations, we have the first-order expression:

$$\delta A = \delta Q R + Q \delta R.$$

Pre-multiplying by $Q_\perp^T$ and post-multiplying by R^{-1}, we have

$$Q_\perp^T (\delta A) R^{-1} = Q_\perp^T \delta Q.$$

Here $Q_\perp^T \delta Q$ represents the part of δQ that lies outside the range space of Q. That is,

$$(I - QQ^T)(Q + \delta Q) = Q_\perp Q_\perp^T \delta Q = Q_\perp Q_\perp^T (\delta A) R^{-1}.$$

Using the fact that the norm of the projector is bounded by one, we have

$$\|(I - QQ^T)\delta Q\| \leq \|\delta A\| \|R^{-1}\| = \|\delta A\| / \sigma_n(A).$$

Therefore,

$$\|\delta \Pi\| \leq \sqrt{2}\kappa(A) \frac{\|\delta A\|}{\|A\|}$$

and so

$$\frac{\|\delta r\|}{\|b\|} \leq \frac{\|\delta b\|}{\|b\|} + \sqrt{2}\kappa(A) \frac{\|\delta A\|}{\|A\|}.$$

From our analysis, though, we have seen that the only part of the perturbation to A that matters is the part that changes the range of A.

[1] Demmel's book goes through this argument, but ends up with a factor of 2 where we have a factor of $\sqrt{2}$; the modest improvement of the constant comes from the observation that if $X, Y \in \mathbb{R}^{m \times n}$ satisfy $X^T Y = 0$, then $\|X + Y\|^2 \leq \|X\|^2 + \|Y\|^2$ via the Pythagorean theorem.