Improving the Numerical Stability of `ivmte`

A. Torgovitsky*

December 15, 2020

1 Motivation

Strange numerical warnings and errors from Gurobi are a recurring problem in `ivmte`. They almost certainly result from the optimization problem being poorly scaled, a point that is supported by the scaling statistics reported by Gurobi. This note contains a proposal that should systematically improve scaling.

I will focus on the "direct" procedure for now, since it is easier. After we check that all of this works for the direct procedure, we can try to apply a similar strategy to the original case with IV–like estimands.

2 Problem Setup

- Assume that the MTRs have the following form:

$$\mathbb{E}[Y(d)|U = u, X = x] \equiv m(d|u, x) = \sum_{k=1}^{K} \theta_k b_k(d|u, x), \tag{1}$$

  where $\theta \equiv [\theta_1, \ldots, \theta_K]'$ are unknown coefficients and b_k are known basis functions.

- Assume that $\theta \in \Theta$ can be represented as $r_{\text{lb}} \leq R\theta \leq r_{\text{ub}}$ for some known constraint matrix R and vector r. (In practice, R and r change on each iteration of the audit procedure, but I will ignore this in the notation.)

- Let $p(x, z) \equiv \mathbb{P}[D = 1|X = x, Z = z]$ and $P \equiv p(X, Z)$ as usual.

*Kenneth C. Griffin Department of Economics, University of Chicago.

- As we know, the basis representation implies that

$$\mathbb{E}[Y|D,X,P] = \sum_{k=1}^{K} \theta_k \left(\frac{D}{P} B_k(1|P,X) + \frac{(1-D)}{1-P} B_k(0|P,X) \right) \equiv \sum_{k=1}^{K} \theta_k B_k, \tag{2}$$

where

$$B_k(0|p,x) \equiv \frac{1}{1-p} \int_p^1 b_k(0|u,x)\, du, \quad B_k(1|p,x) \equiv \frac{1}{p} \int_0^p b_k(1|u,x)\, du \tag{3}$$

$$\text{and} \qquad B_k \equiv B_k(D|P,X) = DB_k(1|P,X) + (1-D)B_k(0|P,X). \tag{4}$$

- Denote the least squares criterion by

$$\hat{Q}(\theta) = \sum_{i=1}^{n} \left(Y_i - \sum_{k=1}^{K} \theta_k B_{ki} \right)^2 \equiv \sum_{i=1}^{n} (Y_i - B_i'\theta)^2 = \|Y - B\theta\|^2, \tag{5}$$

using the usual linear model notation.

3 Rescaling the Least Squares Objective

- The general problem is that $\hat{Q}(\theta)$ is a poorly scaled quadratic form. That is, the matrix $B'B$ has elements (not counting zeros) that are of dramatically different orders.

- This problem is sort of inherent to the MTR problem, because some of the columns of B will be columns for the "u" portions of the MTR, which live between $[0,1]$, and might be squared, cubed, etc., while other columns of B are for the "x" portions, and might be something like year of birth in the AE data, which could be on a dramatically different scale. Relying on the user to fix this problem by scaling their x's is annoying for the user. But more importantly, it doesn't solve the problem for the u portion, and the user cannot get direct access to this portion.

- The idea is to simply rescale each column of B to lie in $[0,1]$ by defining:

$$\tilde{B}_{ki} \equiv \frac{B_{ki} - \text{lb}_k}{(\text{ub}_k - \text{lb}_k)}, \tag{6}$$

where lb_k and ub_k are the minimum and maximum of $\{B_{ki}\}_{i=1}^n$.

- Is it possible for B_k to be constant, so that $\text{ub}_k = \text{lb}_k$ and $\tilde{B}_k$ does not exist? I don't think it is, because of the way $m(0|u, x)$ and $m(1|u, x)$ are specified separately in `ivmte`. So for example if we have a constant term in the MTR for $d = 1$, then $b_k(d|u, x) = d$, so that $B_k = D$. If we also have a constant in the MTR for $d = 0$, this shows up as $B_{k'} = (1 - D)$ for some other index $k' \neq k$.

- Substituting into (5), we can write it as

$$\begin{aligned}
\hat{Q}(\theta) &= \sum_{i=1}^{n} \left(Y_i - \left(\sum_{k=1}^{K} \theta_k (\text{ub}_k - \text{lb}_k) \frac{(B_{ki} - \text{lb}_k)}{(\text{ub}_k - \text{lb}_k)} + \theta_k \text{lb}_k \right) \right)^2 && (7) \\
&= \sum_{i=1}^{n} \left(Y_i - \left(\sum_{k=1}^{K} \theta_k \text{lb}_k \right) - \left(\sum_{k=1}^{K} \theta_k (\text{ub}_k - \text{lb}_k) \tilde{B}_{ki} \right) \right)^2 && (8) \\
&\equiv \sum_{i=1}^{n} \left(Y_i - \xi_0 - \sum_{k=1}^{K} \xi_k \tilde{B}_{ki} \right)^2 \equiv \tilde{Q}(\xi). && (9)
\end{aligned}$$

where

$$\xi_0 \equiv \sum_{k=1}^{K} \theta_k \text{lb}_k \quad \text{and} \quad \xi_k = \theta_k (\text{ub}_k - \text{lb}_k) \text{ for all } k = 1, \ldots, K. \tag{10}$$

- In (9) we now have a least squares criterion where the quadratic form is going to be well-behaved, as all columns of $\tilde{B}$ lie in $[0, 1]$. The cost is that we have a constant term now, and that the coefficients ξ_k are rescaled versions of the parameters we actually want. The first cost is negligible. The second cost means we need to keep track of the difference between ξ_k and θ_k. Essentially the idea is to load all of the scale differences onto the variables of optimization and away from the fixed inputs of the optimization problem.

- We have not had any numerical stability problems in the point identified case because we use `lm`, which works by solving the normal equation after a QR decomposition and is much more stable numerically. So it's probably unnecessary to use the rescaled form (9) for the point identified case. However, *it couldn't hurt*, and—more immediately useful for our purposes—it is an excellent way to debug any issues with rescaling.

4 Optimization

- In the partially identified case, we first want to solve for:

$$\hat{Q}^\star \equiv \min_{\theta \in \mathbb{R}^{d_\theta}} \hat{Q}(\theta) \quad \text{s.t.} \quad r_{\text{lb}} \leq R\theta \leq r_{\text{ub}}. \tag{11}$$

- Now we just want to change variables from θ to ξ. Note that the jth row of $R\theta$ can be written as

$$[R\theta]_j \equiv \sum_{k=1}^{K} R_{jk}\theta_k = 0 \times \xi_0 + \sum_{k=1}^{K} \left(\frac{R_{jk}}{\text{ub}_k - \text{lb}_k} \right) \xi_k \equiv \sum_{k=0}^{K} \tilde{R}_{jk}\xi_k \equiv [\tilde{R}\xi]_j. \tag{12}$$

  Hopefully $\tilde{R}$ is also scaled better, or at least not dramatically worse than R.

- Then instead of (11), we solve

$$\tilde{Q}^\star = \min_{\xi \in \mathbb{R}^{d_\theta + 1}} \tilde{Q}(\xi) \quad \text{s.t.} \quad r_{\text{lb}} \leq \tilde{R}\xi \leq r, \tag{13}$$

  and we should have $\tilde{Q}^\star = \hat{Q}^\star$.

- In the second step we want to solve for

$$\hat{t}_{\text{lb}} \equiv \min_{\theta \in \mathbb{R}^{d_\theta}} \hat{\tau}'\theta \quad \text{s.t.} \quad r_{\text{lb}} \leq R\theta \leq r_{\text{ub}}, \quad \text{and} \quad \hat{Q}(\theta) \leq \hat{Q}^\star(1 + \kappa). \tag{14}$$

- So we need to apply our scaling to the objective here as well:

$$\hat{\tau}'\theta \equiv \sum_{k=1}^{K} \hat{\tau}_k\theta_k = 0 \times \xi_0 + \sum_{k=1}^{K} \left(\frac{\hat{\tau}_k}{\text{ub}_k - \text{lb}_k} \right) \xi_k \equiv \sum_{k=0}^{K} \tilde{\tau}_k\xi_k \equiv \tilde{\tau}'\xi. \tag{15}$$

- Then instead of (11) we solve

$$\tilde{t}_{\text{lb}} \equiv \min_{\xi \in \mathbb{R}^{d_\theta + 1}} \tilde{\tau}'\xi \quad \text{s.t.} \quad r_{\text{lb}} \leq \tilde{R}\xi \leq r_{\text{ub}}, \quad \text{and} \quad \tilde{Q}(\xi) \leq \hat{Q}^\star(1 + \kappa), \tag{16}$$

  and we should have $\tilde{t}_{\text{lb}} = \hat{t}_{\text{lb}}$.