GMMSensitivity: Sensitivity analysis in moment condition models

Michal Kolesár

July 29, 2020

Contents

Introduction

The package `GMMSensitivity` implements estimators and confidence interval for sensitivity analysis in moment condition models considered in Armstrong and Kolesár [2020]. In this vignette, we demonstrate the implementation of these confidence intervals using the demand model for automobiles from Berry et al. [1995].

The package includes the dataset `blp`, which contains estimates of the Berry et al. [1995] model, as implemented by Andrews et al. [2017]. A description of the dataset can be obtained via the documentation in R, using `?blp`. We give additional description here:

`G` Matrix with 31 rows and 17 columns, estimate of derivative of the moment condition $G = \partial E[g(w_i, \theta)]/\partial\theta$, evaluated at initial estimate $\hat{\theta}_{\text{initial}}$. The initial estimate corresponds to a second-step GMM estimate, as computed by Andrews et al. [2017].

`H` Vector of length 17, estimate of derivative of average markup $h(\theta)$, evaluated at $\hat{\theta}_{\text{initial}}$.

`W` Weight matrix used to obtain the estimate $\hat{\theta}_{\text{initial}}$. It corresponds to an esitmate of the inverse of the variance of the moment condition Σ.

`g_init` Average moment condition $n^{-1}\sum_{i=1}^{n} g(w_i, \theta)$, evaluated at estimate $\hat{\theta}_{\text{initial}}$ of θ from Berry et al. [1995].

`h_init` Estimate of the average markup, $h(\hat{\theta}_{\text{initial}})$.

`names` Two lists, one for names of the moment conditions, and one for elements of theta

`ZZ` Gram matrix of the instruments, used to specify the set $\mathcal{C}$

`Sig` Variance of moment condition, estimate of Σ, given by the sample variance of the moment condition evaluated at $\hat{\theta}_{\text{initial}}$.

`sdZ` vector of standard deviations of the instruments

`perturb` scaling parameter to give interpretable meaning to violations of moment conditions. For demand-side moments, it corresponds to an estimate of $\delta_d = 0.01\overline{p}/(\overline{y}/\alpha)$, and for supply-side moments, it corresponds to an estimate of $\delta_s = -0.01\overline{p}/\overline{mc}$, where $\overline{p}$ is average price, $\overline{y}$ is average income, $\overline{mc}$ is average marginal cost, and α is a parameter in the utility function. With this scaling, if a given demand-side instrument enters the utility function with coefficient δ_d, a consumer is willing to pay 1% of the average car price for a unit increase in the instrument. If a given supply-side instrument enters the cost function with coefficient δ_s, increasing the instrument by one unit decreases the marginal cost by 1% of the average car price.

`n` Sample size, number of model/years.

Usage

The package implements estimators, confidence intervals, and efficiency calculations for the model (in the notation of Armstrong and Kolesár [2020])

$$g(\theta_0) = c/\sqrt{n}, \quad c \in \mathcal{C}, \quad \mathcal{C} = \{B\gamma \colon \|\gamma\|_p \leq M\}.$$

Suppose that we want to allow all excluded instruments in the Berry et al. [1995] application to be potentially invalid. Fix B to a scaling matrix so that if the jth supply-side instrument is invalid with $\gamma_{sj} = 1$, this means that changing the instrument by one standard deviation changes the marginal cost by γ_{sj}% of the average car price, and if the j the demand-side instrument is invalid with $\gamma_{dj} = 1$, then the consumer willingness to pay for one standard deviation change in the instrument is γ_{dj}% of the average 1980 car price. Let $p = 2$, and $M = \sqrt{\#I}$, where $\#I$ is the number of invalid instruments so that $\gamma = 1$ is included in the set (this is the same scaling as described in Armstrong and Kolesár [2020] in the paper). Then the confidence interval can be constructed as follows:

```r
## Construct estimate of initial sensitivity
blp$k_init <- -drop(blp$H %*% solve(crossprod(blp$G, blp$W %*%
    blp$G), crossprod(blp$G, blp$W)))
## list collecting initial estimates of H, G, Sigma, n,
## g(thetahat), initial sensitivity k, and initial
## estimate of average markup h(thetahat)
eo <- list(H = blp$H, G = blp$G, Sig = blp$Sig, n = blp$n,
    g_init = blp$g_init, k_init = blp$k_init, h_init = blp$h_init)
## Rows corresponding to invalid instruments
I <- vector(mode = "logical", length = nrow(eo$G))
I[c(6:13, 20:31)] <- TRUE
## Matrix B, scaled as described in the paper
B0 <- blp$ZZ %*% diag(sqrt(blp$n) * abs(blp$perturb)/blp$sdZ)
## Value of M
M0 <- sqrt(sum(I))
## Select columns of B0 corresponding to invalid
## instruments
OptEstimator(eo, B0[, I], M = M0, p = 2, alpha = 0.05, opt.criterion = "FLCI")
#>
#>
#> |Estimate  |Max. bias  |SE         |CI                    |
#> |:---------|:----------|:----------|:---------------------|
#> |0.5598804 |0.06295884 |0.02268748 |(0.459604, 0.6601569) |
```

The efficiency κ_* for this confidence interval can be computed using the `EffBounds` function (which can also be used to compute efficiency of one-sided confidence intervals):

```
EffBounds(eo, B0[, I], M = M0, p = 2)$twosided
#> [1] 0.9704414
```

In contrast the CI based on the initial estimate is much wider:

```
OptEstimator(eo, B0[, I], M = M0, p = 2, alpha = 0.05, opt.criterion = "Valid")
#>
#>
#> |Estimate  |Max. bias |SE         |CI                       |
#> |:---------|:---------|:----------|:------------------------|
#> |0.3271788 |0.1983659 |0.01815665 |(0.09894796, 0.5554097) |
```

A specification test for whether the value $M = M0$ is too low, that is a test of the hypothesis $H_0\colon M \leq M0$, can be conducted using the `Jtest` function:

```
## Update eo so that Sig corresponds to the initial
## estimate of Sigma, so that thetahat_initial minimizes
## n*g(theta)Sig^{-1} g(theta), and the J statistic is
## given by the value of this minimum.
eoJ <- eo
eoJ$Sig <- solve(blp$W)
jt <- Jtest(eoJ, B0[, I], M = M0, p = 2, alpha = 0.05)
```

Here `J` is the J-statistic, `p0` is the p-value of the usual J-test (that assumes $M = 0$ under the null), `pC` is the p-value of the test, and `Mmin` is the smallest value of M that is not rejected. Rescaling by $\sqrt{\#I}$, we obtain the result in Table 1 in Armstrong and Kolesár [2020]:

```
jt$Mmin/M0
#> [1] 1.130813
```

If were only concerned about the validity of the demand-side instrument `"demand_firm_const"` (number of cars produced by the same firm), then since this is the sixth instrument, an analogous analysis could be conducted as:

```
I <- vector(mode = "logical", length = nrow(eo$G))
I[6] <- TRUE
OptEstimator(eo, B0[, I, drop = FALSE], M = 1, p = 2, alpha = 0.05,
    opt.criterion = "FLCI")
#>
#>
#> |Estimate  |Max. bias   |SE         |CI                      |
#> |:---------|:-----------|:----------|:-----------------------|
#> |0.3564058 |0.002489518 |0.01868669 |(0.3194578, 0.3933538) |
EffBounds(eo, B0[, I, drop = FALSE], M = 1, p = 2)$twosided
#> [1] 0.8593365
OptEstimator(eo, B0[, I, drop = FALSE], M = 1, p = 2, alpha = 0.05,
    opt.criterion = "Valid")
#>
#>
#> |Estimate  |Max. bias  |SE         |CI                       |
```

```
#> |:---------|:----------|:----------|:----------------------|
#> |0.3271788 |0.01070908 |0.01815665 |(0.2862084, 0.3681493) |
Jtest(eoJ, B0[, I, drop = FALSE], M = 1, p = 2, alpha = 0.05)
#> $J
#> [1] 426.7276
#>
#> $p0
#> [1] 0
#>
#> $pC
#> [1] 0
#>
#> $Mmin
#> [1] 10.20545
```

References

Isaiah Andrews, Matthew Gentzkow, and Jesse M. Shapiro. Measuring the sensitivity of parameter estimates to sample statistics. *Quarterly Journal of Economics*, 132(4):1553–1592, November 2017.

Tim Armstrong and Michal Kolesár. Sensitivity analysis using approximate moment condition models. arXiv:1808.07387, July 2020.

Steven T. Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890, 1995.