ECON 525: Homework 5

Shabab Ahmed

16 February , 2022

1 Problem 1:

Let $\phi(x)$ be the pdf of the standard normal distribution.

(a) Compute $\int_{-1.96}^{1.96} \phi(x)\, dx$ using the Gauss-Chebyshev Quadrature with various numbers of evaluation points.

Solution:

I computed $\int_{-1.96}^{1.96} \phi(x)\, dx$ using various numbers of evaluation points. Specifically, I used $n = [10, 20, 30, 50, 75, 100, 1000, 10000, 50000]$ evaluation points. The computational result using each of these numbers of evaluation points is given below:

n	$\int_{-1.96}^{1.96} \phi(x)\, dx$
10	0.950918520278555
20	0.950237959829586
30	0.950108534888065
50	0.950041847632036
75	0.950020948930560
100	0.950013627737087
1000	0.950004303912142
10000	0.950004210645651
50000	0.950004209741242

As we increase the number of evaluation points, we get greater precision. However, even with 50 evaluation points we get precision upto the 4th decimal place. Therefore, there is only a small marginal gain in increasing the number of evaluation points beyond 50.

(b) Compute $\int_{-1.96}^{1.96} \phi(x)\, dx$ using a Monte Carlo integration.

Solution:

I used $n = 10,000$ evaluation points. Using the method described in the lecture notes, I drew 10,000 evaluation points from the uniform distribution on $[-1.96, 1.96]$. The Monte Carlo integration gives us the following computational result:

$$\int_{-1.96}^{1.96} \phi(x)\, dx \approx \frac{1}{n}\sum_{i=1}^{n} \frac{f(x_i)}{g(x_i)} = \frac{1.96 + 1.96}{n} \times \sum_{i=1}^{n} f(x_i) = \boxed{0.9494087003}.$$

(c) Compare these results with the exact value.

Solution:

Notice: $\int_{-1.96}^{1.96} \phi(x)\, dx = \Phi(1.96) - \Phi(-1.96)$ where Φ is the standard normal CDF. Therefore, the exact value of $\int_{-1.96}^{1.96} \phi(x)\, dx$ is 0.950004209703559. We compute the differences for the two different methods below:

Gauss-Chebyshev Quadrature:

n	$\int_{-1.96}^{1.96} \phi(x)\, dx$
10	9.1431×10^{-4}
20	2.3375×10^{-4}
30	1.0433×10^{-4}
50	3.7638×10^{-5}
75	1.6739×10^{-5}
100	9.4180×10^{-6}
1000	9.4209×10^{-8}
10000	9.4209×10^{-10}
50000	3.7683×10^{-11}

We can see that the difference with the exact value decreases as we increase the number of evaluation points. This should intuitively make sense because as we increase the number of evaluation points we are getting closer to a continuous interval and so the approximation becomes more precise. We can also see that if our tolerance level was 10^{-5} we would achieve the approximation with only 50-75 evaluation points.

Monte Carlo Integration:

The difference with the exact value using Monte Carlo integration is 5.9551×10^{-4}.We can see that with the same number of evaluation points the Gauss-Chebyshev Quadrature performs exceedingly better than Monte Carlo integration. This probably has to do with the randomness in the Monte Carlo integration. Unlike Gauss-Chebyshev quadrature, the points selected in this method does not follow a particular curve. It is possible that the random draw picks the same point several times and this would not provide us with any new information. We also cannot make a direct claim that the approximation will get strictly better as we increase the number of evaluation points due to the randomness associated with it. If given a choice, we should be preferring the Gauss-Chebyshev quadrature method for this example.

---

2 Problem 2:

Consider the following binary-choice model. The two choices are denoted by $\{0, 1\}$. The indirect utility of agent i from choosing these options is given by

$$\begin{aligned} u_{i1} &= \alpha + \beta x_i + \epsilon_{i1} \\ u_{i0} &= 0, \end{aligned}$$

where ϵ_{i1} is the unobserved (to the econometrician) random shock that follows the standard normal distribution and x_i is an exogenous covariate that affects agent's decision. Agent i's optimal decision is

$$d_i = \mathbb{1}(\alpha + \beta x_i + \epsilon_{i1} > 0).$$

We observe $(d_i, x_i)_{i=1}^N$ in the data and our goal is to estimate (α, β).

data.csv contains the random sample (d_i, x_i) for $N = 5,000$ individuals. You will estimate the parameters and their standard errors. First, estimate (α, β) using MLE. Next, compute the asymptotic variance-covariance matrix of the estimators. To do so, use the inverse of outer product of scores. To compute scores, use

---

Solution: (MLE)

Suppose y_i^* is an unobservable (latent) variable such that

$$y_i^* = \alpha + \beta x_i + \epsilon_{i1}$$

We observe $d_i = 1$ if $y_i^* > 0$ and $d_i = 0$ if $y_i^* \geq 0$. Therefore,

$$\begin{aligned} P(d_i = 1|x_i) &= P(y_i^* > 0|x_i) = P(\alpha + \beta x_i + \epsilon_{i1} > 0) \\ &= P(\epsilon_{i1} > -\alpha - \beta x_i) = 1 - \Phi(-\alpha - \beta x_i) \\ &= \Phi(\alpha + \beta x_i) \end{aligned}$$

where $\Phi(\cdot)$ is the CDF of the standard normal distribution. Similarly, we can show that

$$P(d_i = 0|x_i) = 1 - \Phi(\alpha + \beta x_i)$$

Thus, the likelihood of observing datapoint i is going to be given by:

$$L_i(\alpha, \beta) = \Phi(\alpha + \beta x_i)^{d_i} \times (1 - \Phi(\alpha + \beta x_i))^{1-d_i}$$

Using independence, the likelihood of observing the dataset is given by:

$$L(\alpha, \beta) = \prod_{i=1}^{n} L_i(\alpha, \beta)$$

Taking logs:

$$\ln\left(L(\alpha, \beta)\right) = \sum_{i=1}^{n} d_i \ln\left(\Phi(\alpha + \beta x_i)\right) + (1 - d_i) \ln\left(1 - \Phi(\alpha + \beta x_i)\right)$$

The maximum likelihood estimate is then given by:

$$(\widehat{\alpha}, \widehat{\beta}) = \arg\max_{\alpha, \beta} \sum_{i=1}^{n} d_i \ln\left(\Phi(\alpha + \beta x_i)\right) + (1 - d_i) \ln\left(1 - \Phi(\alpha + \beta x_i)\right)$$

For our purposes, we minimize the negative log likelihood. We use MATLAB's *fminsearch* function with an initial guess of $(\alpha, \beta) = (1, 1)$ to obtain the following result:

$$\begin{bmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{bmatrix} = \begin{bmatrix} 1.6351 \\ -0.3006 \end{bmatrix}$$

(a) the first difference formula (with the varying step size)

Solution:

The score vector (or score) is the gradient of the log-likelihood function with respect to the parameters being estimated. Let $F(\alpha, \beta) = \ln\left(L(\alpha, \beta)\right)$. The score is given by:

$$\nabla F = \begin{bmatrix} \frac{\partial F}{\partial \alpha} \\ \frac{\partial F}{\partial \beta} \end{bmatrix}$$

We need to approximate each of the partial derivatives to approximate the score. The forward difference formula gives us the following:

$$\frac{\partial F}{\partial \alpha} \approx \frac{F(\alpha + h, \beta) - F(\alpha, \beta)}{h}$$
$$\frac{\partial F}{\partial \beta} \approx \frac{F(\alpha, \beta + h) - F(\alpha, \beta)}{h}$$

The above expressions will be evaluated at the estimated parameter values, $\widehat{\alpha}$ and $\widehat{\beta}$. We calculate the variance covariance matrix using the inverse of the outer product of scores. We approximate the standard errors using a range of h values to see which value gives us the best approximation. In particular, we consider the following h values: $\{1, 1e-1, 1e-2, 1e-3, 1e-4, 1e-5, 1e-6\}$. The following table contains the standard error approximations for the various step sizes along with the differences with the standard errors computed using the closed form. Finally, the table also includes the norm of the difference between the two standard errors. The reason we add the last two pieces of information is because the standard error approximations are identical upto four decimal points for certain values of h. Therefore, we use the difference to measure the performance of a particular step size. Moreover, we use the norm because we have a tuple and we want to make sure that we have a measure of differences in standard error approximations of both estimates. We report a tuple for the standard error approximation where the first element corresponds to the standard error of $\widehat{\alpha}$ and the second element corresponds to the standard error of $\widehat{\beta}$

h	Standard errors	Difference	Norm difference
1	[0.0290, 0.0037]	[0.0253, 0.0103]	0.0273
1e-1	[0.0512, 0.0129]	[0.0031, 0.0011]	0.0033
1e-2	[0.0543, 0.0140]	[4.1276e-5, 1.1106e-5]	4.2744e-5
1e-3	[0.0543, 0.0140]	[8.1778e-6, 2.4802e-6]	8.5456e-6
1e-4	[0.0543, 0.0140]	[8.5815e-7, 2.6168e-7]	8.9716e-7
1e-5	[0.0543, 0.0140]	[8.6218 e-8, 2.6304e-8]	9.0141e-8
1e-6	[0.0543, 0.0140]	[8.6281e-9, 2.6317e-9]	9.0206e-9

Interestingly, in our case we find that the smaller step size we take the better our approximation gets to the closed form standard errors. This can be seen by looking at the differences and also at the norm of the difference. It also seems like the difference is decreasing by a factor of decrease in h. I suspect this result might not hold if we keep decreasing h due to competing roundoff errors. In fact, I checked for $h = 1e - 10$ and found that the norm difference is actually greater than that of $h = 1e - 6$.

(b) the central difference formula (with the varying step size)

Solution:

The central difference formula gives us the following approximations of the partial derivatives:

$$\frac{\partial F}{\partial \alpha} \approx \frac{F(\alpha + h, \beta) - F(\alpha - h, \beta)}{2h}$$
$$\frac{\partial F}{\partial \beta} \approx \frac{F(\alpha, \beta + h) - F(\alpha, \beta - h)}{2h}$$

We do a similar analysis as we did for first difference formula. We use $h \in \{1, 1e-1, 1e-2, 1e-3, 1e-4, 1e-5, 1e-6\}$ and report the approximate standard errors, the difference with closed form standard errors and the norm of the differences. The following table contains these results:

h	Standard errors	Difference	Norm difference
1	[0.0381, 0.0057]	[0.0162, 0.0083]	0.0182
1e-1	[0.0538, 0.0137]	[5.0096e-4, 2.5932e-4]	5.6410e-4
1e-2	[0.0543, 0.0140]	[5.0863e-6, 2.6369e-6]	5.7292e-6
1e-3	[0.0543, 0.0140]	[5.0871e-8, 2.6374e-8]	5.7301e-8
1e-4	[0.0543, 0.0140]	[5.0872e-10, 2.6374e-10]	5.7302e-10
1e-5	[0.0543, 0.0140]	[5.5026e-12, 2.6596e-12]	6.1117e-12
1e-6	[0.0543, 0.0140]	[4.1393e-12, 2.8847e-13]	4.1493e-12

We observe a similar phenomenon as to the forward difference case. We see that the difference and the norm of the difference is decreasing for our range of h values. However, in this case the decrease in difference seems to be a factor of the square of the decrease in step size. Again, we do not expect this pattern to hold for smaller values of h due to rounding errors. In fact, we checked for $h = 1e - 10$ and found that difference and norm difference is higher than $h = 1e - 6$. However, we can note that for any given h the central difference formula does better than the forward difference formula. This intuitively makes sense because the central difference formula is using more information in a way.

(c) the closed form.

Solution:

We can calculate the closed form of the partial derivatives of the objective function with respect to the parameters. Recall the score is given by:

$$\nabla F = \begin{bmatrix} \frac{\partial F}{\partial \alpha} \\ \frac{\partial F}{\partial \beta} \end{bmatrix}$$

We can calculate the following:

$$\begin{aligned} \frac{\partial F}{\partial \alpha} &= \sum_{i=1}^{n} \frac{d_i \phi(\alpha + \beta x_i)}{\Phi(\alpha + \beta x_i)} - \frac{(1 - d_i)\phi(\alpha + \beta x_i)}{1 - \Phi(\alpha + \beta x_i)} \\ \frac{\partial F}{\partial \beta} &= \sum_{i=1}^{n} \frac{x_i d_i \phi(\alpha + \beta x_i)}{\Phi(\alpha + \beta x_i)} - \frac{x_i(1 - d_i)\phi(\alpha + \beta x_i)}{1 - \Phi(\alpha + \beta x_i)} \end{aligned}$$

The above expressions will be evaluated at our estimated parameter values, $\widehat{\alpha}$ and $\widehat{\beta}$. Therefore, we can calculate the above expressions to get the score. Then, using the inverse of the outer product of scores we get the following variance covariance matrix:

$$V = \begin{bmatrix} 0.0029 & -7.0433e - 04 \\ -7.0433e - 04 & 1.9512e - 04 \end{bmatrix}$$

The square root of the diagonal elements of the variance covariance matrix gives us the standard errors. The following table contains the estimates along with their standard errors.

Parameter	Estimate	Standard error
α	1.6351	0.0543
β	-0.3006	0.0140