Applied Macroeconomics F, Lecture Note
Integration

Naohito Abe

2020 Summer

1 Introduction

In many applied economic models, we need to conduct "integration". For example, in continuous time dynamic models such as Ramsey's optimal growth model, the utility function is defined as an integral over time. When aggregating the entire economy over continuously many agents, the aggregate consumption or income is defined as an integral over agents. But most frequently, integration becomes an important when we include stochastic factor in our model. The Euler equation for intertemporal optimization becomes an integral over a random variable. To obtain the permanent income, we also need integrations if there are any uncertainty in future income. The maximum likelihood estimators is nothing but the optimization of an integral. Whenever we want to solve an economic model with uncertainty or heterogeneity, some forms of integration is necessary.

2 Simple Numerical Integration

Numerical integration is much easier to implement than numerical differentiation. A very intuitive method is the midpoint estimator, which is easy to understand because it is very similar to the definition of Riemann-Integral. Basically, the midpoint estimator is an approximation of a venation by a rectangular. Take a look at Figure 1. We are interested in an integration of a function $f(x)$ from a to b. The value of the integration is illustrated as the area between function $f(x)$ and the x-axis. Suppose we know that the function passes points A, C, and D. C is $((b-a)/2, f((b-a)/2))$,that is, y axis of C is the value of $f(x)$ at the midpoint of the interval between a and b. One candidate of an approximation of the integral can be given by a square, $a-P-R-b$, which is an approximation of the area by a square. Since the height of $f(x)$ is evaluated at the midpoint of the interval, this approximation is called as the midpoint rule. The sum of two trapezoids, $a-A-C-(b-1)/2$, and $(b-a)/2-C-D-b$ can be also regarded as an approximation of the area, which is the approximation by a line AC, and CD. This approximation is sometimes called as the trapezoid

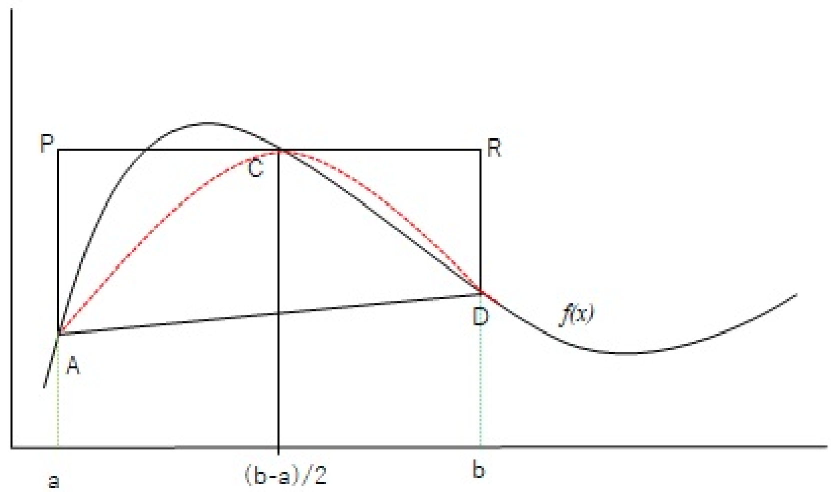

Figure 1: Midpoint Rule and Newton-Cotes

rule.[1] Finally, we can think of an area below the parabola (quadratic function) $A - C - D$. This approximation uses information of three points, A, C,and D. Then, we approximate the function by a polynomial. This method is called as Newton-Cotes quadrature[2].

Although these rules are easy to understand, these are seldom used in applied economics. As the number of points increases, the integration tends to converge to the true value (no always, though). The reason is simple. They are slow or fail to converge to the true integral. See Judd (1998) for detailed discussion and an example of the slowness of these rules when we try to integrate x^{-2} from 1 to 10. Rather, most people use the Gaussian quadrature or the (Quasi) Monte Carlo methods. The Gaussian quadrature is much more accurate than the Newton-Cotes or the Simpson's method. It is reasonably fast in modern PCs. For integrations under higher dimensions, people have begun to use the Monte Carlo method.[3]

A word "quadrature" has several meanings. According to the Oxford dictionary, quadrature means "The process of constructing a square with an area equal to that of a circle, or of another figure bounded by a curve with compass and straightedge" Of course it is impossible to draw a square with the same

[1] There is a variant of the trapezoid rule, called the Simpson's rule, which is discussed in many textbooks on numerical methods such as Judd (1998).

[2] The term "quadrature" originally means an ancient quiz in Greek to find a process to construct a rectangular (or finite set rectangular) that has the same area of a given figure such as a circle with compass and straightedge. Of course, it is impossible to

[3] See Train (2009) *Discrete Choice Methods with Simulation* for the detail of simulation based integration.

area of a given circle with only a finite step. But, it is not difficult to get a good approximation. By increasing the number of steps, the approximation becomes very accurate. This idea is very close to the original Riemann-Integral. Now, a word quadrature is used in more general numerical procedures for integration.

3 Basic Idea of Gaussian Quadrature

We are interested in finding,

$$\int_a^b f(x)\, dx, \tag{1}$$

for given interval $[a, b]$ and a known function, f. It is well known that except for some special cases, we cannot get the integral as a function of a and b in a closed form. The idea of the Gaussian quadrature is to find an approximation by a sum of polynomials that are exact when the function f is actually a polynomial of some degree of $(\leq m)$. You might notice that Newton-Cotes also uses a polynomial. The crucial difference between Gaussian quadrature and Newton-Cotes is that the former uses the orthogonal polynomials, just like we discussed in Cheybyshev polynomial while the latter uses the second order natural polynomial only.

To implement Gaussian quadrature, given n, the number of grids to approximate the integral, we would like to find a combination of $(\alpha_1, \alpha_2, ..., \alpha_n)$, and $(x_1, x_2, ..., x_n)$ such that

$$\int_a^b f(x)\, w(x)\, dx \approx \sum_{i=1}^{n} \alpha_i f(x_i), \tag{2}$$

$$\int_a^b f(x)\, w(x)\, dx = \sum_{i=1}^{n} \alpha_i f(x_i) \text{ if } f \text{ is polynomials of degree } \leq m, \tag{3}$$

where $w(x)$ is a weighting function.

Following Judd (1998), we formulate a famous Theorem by Gauss (1816).[4]

Let a closed interval $[a, b]$ and the number $n \in R$ be fixed. Suppose that $\{\varphi_i(x)\}_{i=0}^{\infty}$ is an orthonormal family of polynomials with respect to a weighting function, $w(x)$. Define q_i so that $\varphi_i(x) = q_i x^i + ...$ Let $x_i, i = 1, 2, .., n$ be the n zeros of $\varphi_n(x)$. Then, $a < x_1 < x_2 < ... < x_n < b$, and if f is twice continuously differentiable in $[a, b]$,then, $\int_a^b f(x)\, w(x)\, dx \approx \sum_{i=1}^{n} \alpha_i f(x_i)$.

[4] Usually, I told students that we should look for the original paper when using theorems, or citing results. However, in mathematics, many important works are written in Latin. This quite famous paper by Gauss is no exception, which makes it very diffiucult for us to look for the originals. Some very important theorems or results in economic theory are also written in other than English, such as many papers by Walras (in French), Pareto (in Italian), and Laspeyres (in German). When I actually read papers by Laspeyes in German (with helps of translation by machines), it was a big surprise. The objects, the way of discussion, and the main massages are so different from those we imagined from standard textbooks in Macroeconomics. Anyway, it is worth reading "classics" if possible.

The weights for each point, x_i is given by

$$\alpha_i = \frac{-q_{n+1}/q_n}{\varphi_n'(x_i)\,\varphi_{n+1}(x_i)} > 0$$

We do not define the approximation, $\approx$, above, thus the above is not rigorous. However, we can see that we can find $(\alpha_1, \alpha_2, ..., \alpha_n)$, and $(x_1, x_2, ..., x_n)$, it should be straightforward to take an integral of any twice continuously differentiable functions, f. The issue here is how to find an actual set of $(\alpha_1, \alpha_2, ..., \alpha_n)$, and $(x_1, x_2, ..., x_n)$. As the above theorem says, it depends on what kinds of polynomials we use for (3).

4 Gauss-Legendre Quadrature

Gauss-Legendre quadrature is often used in economics. Suppose we are interested in the following integration,

$$\int_{-1}^{1} f(x)\,dx. \tag{4}$$

Note that it is easy to change the interval from $[-1, 1]$ to $[a, b]$ without changing the main results. Different quadratures are used when we use "weights" such as,

$$\int_{0}^{\infty} f(x)\exp(-x)\,dx : \text{ Gauss-Laguerre,} \tag{5}$$

$$\int_{-\infty}^{\infty} f(x)\exp\left(-x^2\right)dx : \text{ Gauss-Hermite,} \tag{6}$$

$$\int_{-1}^{1} f(x)\left(1-x^2\right)^{-1/2} dx : \text{ Gauss-Chebyshev,} \tag{7}$$

$$\int_{-1}^{1} f(x)(1+x)^a (1-x)^b\,dx : \text{ Gauss-Jacobi.} \tag{8}$$

In many casews, when we are interested in integrations, we do not have such weights. Even if there are weights, by incorporating the weights into the function, f, we can eliminate the weight. Therefore, Gauss-Legendre quadrature is a natural candidate for our integration. However, by transforming the variable, x , it is possible to apply other integrations such as Gauss-Chebyshev. However, I don't think there are much advantages in doing so. The name "Gauss-Legendre quadrature" comes from the fact that we use the Legendre polynomials for (3). Note that similar to the Chebyshev polynomials, the Legendre polynomials are orthonormal, which gives us a very nice base to approximate functions.

Actual procedure to get $(\alpha_1, \alpha_2, ..., \alpha_n)$, and $(x_1, x_2, ..., x_n)$ is a bit cumbersome. See Mori, Murota, and Sugihara (1992) for example. Compecon toolbox provides us with a function that gives $(\alpha_1, \alpha_2, ..., \alpha_n)$, and $(x_1, x_2, ..., x_n)$.

Note that when we choose the Legendre polynomials for our integration, the points for evaluation, $(x_1, x_2, ..., x_n)$ and their weights$(\alpha_1, \alpha_2, ..., \alpha_n)$, do not depend on the actual function, f. When I first encountered this idea of the Gaussian quadrature, this independence seemed very odd. For example, a famous software, *Mathematica*, conducts a numerical integration in two steps, first it plots f roughly to detect the regions in which f changes a lot. Then, the software increases the number of grid assigned to the regions. In other words, the choice of $(\alpha_1, \alpha_2, ..., \alpha_n)$, and $(x_1, x_2, ..., x_n)$ depends on the shape of f, which sounds very natural. The trick of Gaussian quadrature is it takes advantages of approximation by orthonormal polynomials. Any functions can be approximated by polynomials. It becomes a good approximation if the function well behaves. So, if the number of quadrature point, $(x_1, x_2, ..., x_n)$, is extremely small, such as 2 or 3, the performance of the Gaussian quadrature is not so good. However, if the number exceeds 10, the Gaussian quadrature becomes very accurate because the Legendre polynomials become a very good approximation of the original function,f.

5 Matlab Example

The compecon toolbox provided by Miranda and Fackler contains very useful matlab codes to find a set of $(\alpha_1, \alpha_2, ..., \alpha_n)$, and $(x_1, x_2, ..., x_n)$. The below is an example of calculate,

$$\int_0^1 \left(1 + x^2\right) f(x)\, dx, \tag{9}$$

$$f : \text{pdf for lognormal.} \tag{10}$$

Integrating over lognormal distribution is quite popular among consumption research.

To implement the following code, we need both compecon toolbox and statistical toolboxes. The compecon toolbox is needed to obtain the quadrature points, $(x_1, x_2, ..., x_n)$ and the weights,$(\alpha_1, \alpha_2, ..., \alpha_n)$. Statistical toolbox is used for pdf of the lognormal distribution.

```
% Numerical Integration
%
% Integral(1+(x^2 ))f(x);
% f: lognormal with u=0, sig=0.1
clear
n=30;
a=0;
b=1;
%
[x2,w]=qnwlege(n,a,b); % Gauss-Legendre quadrature
```

```
forquadsum=0;
for i=1:n
    forsum1d=w(i)*(1+x2(i)^2) *lognpdf(x2(i),0,0.1);
    forquadsum=forquadsum+forsum1d;
end
%%%%%%
```

6 Monte-Carlo Integration

When we need to integrate over three or more variables, or when the integration includes complicated conditional distribution, it is often the case that Gaussian quadrature does not provide us with nice approximations. In such a case, Monte-Carlo and its variants are the ways you should go. Due to a rapid progress in microeconometrics and Bayesian econometrics, Monte-Carlo integration is increasingly getting popular among researchers. Accordingly, there are several excellent textbooks on the method such as Train (2009). If your integration consists of more than three variables, I strongly recommend that you should read such textbooks. In this note, I briefly introduce Monte-Carlo integration.

Suppose we are interested in an integration of $f(x)$ over $x \in X$.

Suppose we have a set of random draw from X, $(x_1, x_2, .., x_n)$ and $(f(x_1), f(x_x), ..., f(x_n))$. Then, by the strong law of the large number, we must have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} f(x_n) = E[f(x)] \text{ with probability one.} \tag{11}$$

For finite sample, we can expect

$$\frac{1}{n} \sum_{i=1}^{n} f(x_n) \approx E[f(x)]$$

The problem in the above method is that we need a random draw, which is actually very difficult. Of course, most standard package such as R, Matlab, Stata, and Python are able to generate random draw from various distributions. However, most cases, even if they state that they are random draws, most likely, they use deterministic routines. Therefore, most random numbers generated by standard packages are not random, but pseudorandom numbers. For example, whenever you restart Matlab, Matlab returns the identical results for its random generators, rand, randi, and randn. Since randomization is the key for cryptography and one of the central topics in the recent network system, generating truly random draw is a very active research area. When we need very accurate random draw, we need to use a physical random number generator. The basic idea is that we actually observe physical movements that are supposed to be random, such as a dice. Of course, modern physical random generators use more sophisticated system such as noises observed in PC.

To apply (11), we need a random draw. Suppose we do not have a random draw, but have many observations in X that satisfies the following equation,

$$\lim_{n \to \infty} \frac{b-a}{n} \sum_{i=1}^{n} f\left(x_n\right) = \int_a^b f\left(x\right) dx.$$

The basic idea is as follows. When integrating a function, we need a random draw to cover the area evenly. For example, if we are interested in the integration of a function $f(x)$ over [0,1], grids with even intervals must suffice for the purpose.

Recent progresses in Monte-Carlo and its variants are quite rapid, and have been used intensively in Bayesian econometrics such as NKDSGE, which is beyond the scope of this lecture note.