Stochastic Process Fundamentals

Fundamental concepts regarding Sequences, Convergence and Times Series

2019DMB02

0.1 Sequences

A sequence is nothing but a list of numbers written in a specific order. Sequences can be infinite or finite. A general form of sequences can be shown:

a_1 - *first term*
a_2 - *second term*
a_n - n^{th} *term*

Some of the common ways we can denote sequences are as follows:

$$\{a_1, a_2, \cdots, a_n, a_{n+1}, \cdots\} \tag{1}$$

$$\{a_n\} \tag{2}$$

$$\{a_n\}_{n=1}^{\infty} \tag{3}$$

To illustrate with an example, here is how we would write the first few terms of a sequence:

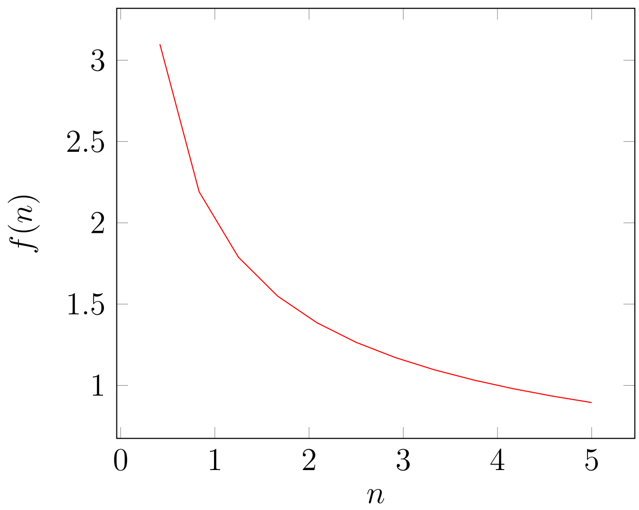
$$\{\frac{n+1}{n^2}\}_{n=1}^{\infty} = \{\underset{n=1}{2}, \underset{n=2}{\frac{3}{4}}, \underset{n=3}{\frac{4}{9}}, \underset{n=4}{\frac{5}{16}}\} \tag{4}$$

An interesting way to think about sequences is as functions that map index values to the value that the particular sequence might take. For example consider the same sequence as above written as a function and its values written in a tuple of the format $(n, f(n))$.

$$f(n) = \frac{n+1}{n^2} \tag{5}$$

$$values \rightarrow (1, 2), (2, 3/4), (3, 4/9), (4, 5/16) \tag{6}$$

We do this because in this situation we can essentially plot out the values and obtain a graphical representation of a sequence.



We can observe from this graph that as n increases the value of sequence terms is going closer and closer to zero. Hence we can say that the limiting value of this sequenec is zero:

$$\lim_{n \to \infty} a_n = \lim_{n \to \infty} \frac{n+1}{n^2} = 0 \tag{7}$$

0.1.1 General theorems and statements

- We can generally state that if we make a_n sufficiently close to L for large values of n then the values of a_n approaches L as n approaches infinity.

  $$\lim_{n \to \infty} a_n = L$$

  As a more precise definition we can say that $\lim_{n \to \infty} a_n = L$ if for every number $\epsilon > 0$ there exists an integer N such that:

  $$|a_n - L| < \epsilon, \textit{when: } n > N$$

- We can say that $\lim_{n \to \infty} a_n = \infty$ if for every number $M > 0$ there is an integer N such that:

  $$a_n > M \textit{ when: } n > N$$

- We can say that $\lim_{n \to \infty} a_n = -\infty$ if for every number $M < 0$ there exists a number N such that:

  $$a_n < M \textit{ when: } n > N$$

- The key insight for us that for a limit to exist and have a finite value, then all the sequence terms must get closer and closer to that finite value as n approaches infinity.

- If $\lim_{n\to\infty} a_n$ exists and is finite we say that the sequence is **convergent** whereas if $\lim_{n\to\infty} a_n$ does not exist and if infinite then we say that the sequence is **divergent**.
- Given a sequence $\{a_n\}$ if we have a function $f(x)$ such that $f(n) = a_n$ and that $\lim_{x\to\infty} f(x) = L$ then we can say that:

$$\lim_{n\to\infty} a_n = L$$

0.1.2 Squeeze theorem

We can state the squeeze theorem for sequences as follows:

if: $a_n \leq c_n \leq b_n$ *for all* $n > N$ *for some* N
and if: $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = L$
Then we can say that: $\lim_{n\to\infty} c_n = L$

This theorem is particularly useful when we are trying to compute the limits of sequence that alternate in signs, for example the modulus function. Another important theorem to state, which we shall prove using the squeeze theorem is:

if: $\lim_{n\to\infty} |a_n| = 0$ *then:* $\lim_{n\to\infty} a_n = 0$

Additionally we note that for this theorem to work, the limit has to be zero. Now to prove this using the squeeze theorem:

We can first of all note that: $-|a_n| \leq a_n \leq |a_n|$
Then we note that: $\lim_{n\to\infty}(-|a_n|) = -\lim_{n\to\infty} |a_n| = 0$
Therefore now that we have: $\lim_{n\to\infty}(-|a_n|) = \lim_{n\to\infty} |a_n| = 0$
then by squeeze theorem we have: $\lim_{n\to\infty} a_n = 0$

As an additional theorem of convergence that is closely related we can state:

The sequence: $\{r^n\}_{n=0}^{\infty}$ *converges if:* $-1 < r < 1$
and diverges for all other values of r

Mathematically this can also be stated as :

$$\lim_{n\to\infty} r^n = \begin{cases} 0, & \text{if } -1 < r < 1. \\ 1, & \text{if } r = 1. \\ \infty & \text{otherwise} \end{cases} \tag{8}$$

0.1.3 Increasing, Decreasing and Bounded

Given a sequence $\{a_n\}$ we have the following important definitions that explain key concepts about the nature of the sequence.

- A sequence is **increasing** if: $a_n < a_{n+1}$ for every n.
- A sequence is **decreasing** if: $a_n > a_{n+1}$ for every n.

- If a_n is an increasing or decreasing sequence it is known to be **monotonic**. Note that a monotonic sequence always either increases or decreases, not both.
- If there exists a number m such that $m \leq a_n$ for every n then we say that the sequence is **bounded below** and m is called the **lower bound** of the sequence.
- If there exists a number M such that $a_n \leq M$ for every n then we say that the sequence is **bounded above** and M is called the **upper bound** of the sequence.
- Finally we can say that if $\{a_n\}$ is bounded and monotonic then $\{a_n\}$ is convergent.

0.2 Series

To begin defining an **infinite series** we first start with a sequence $\{a_n\}$. Note that a sequence is just a sequence of numbers whereas a series represents some kind of operation on those sequence of numbers. We can define a basic series as:

$$\begin{gathered} s_1 = a_1 \\ s_2 = a_1 + s_2 \\ s_3 = a_1 + a_2 + a_3 \\ s_n = \textstyle\sum_{i=1}^{n} a_i \end{gathered}$$

We can further note that the successive values of the series itself forms a sequence of numbers which can be represented as $\{s_n\}_{n=1}^{\infty}$. This is a sequence of **partial sums**. Now we can compute the limiting value of this sequence of partial sums as:

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{i=1}^{n} a_i = \sum_{i=1}^{\infty} a_i \tag{9}$$

Note that as in the case of sequences before, if the sequence of series values has a finite limit, then the series is said to be **convergent** and if the limit does not exist then it is **divergent**. Now we will prove the following theorem :

if: $\sum a_n$ *converges then:* $\lim_{n\to\infty} a_n = 0$

- Step 1: We can write the following two partial sums for the given series:

$$\begin{gathered} s_{n-1} = \textstyle\sum_{i=1}^{n-1} a_i = a_1 + a_2 + \cdots + a_{n-1} \\ s_n = \textstyle\sum_{i=1}^{n} a_i = a_1 + a_2 + \cdots + a_n \end{gathered}$$

- Step 2: Subtracting the two partial sums we can get :

$$a_n = s_n - s_{n-1}$$

- We can say that if $\sum a_n$ is convergent then the sequence of partial sums is also convergent for some finite value. Note that the same holds true for the partial sums series of n and $(n-1)$.

$$\{s_n\}_{n=1}^{\infty} \to \lim_{n\to\infty} s_n = s \to \lim_{n\to\infty} s_{n-1} = s$$

- Step 4: Finally we can write:

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty}(s_n - s_{n-1}) = s - s = 0$$

0.2.1 The Ratio and Root test

The **ratio test** can be applied to check for convergence of a series. Suppose we have a series given by:

$$\sum a_n \tag{10}$$

Then we can define:

$$L = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| \tag{11}$$

Now the following conditions would hold:

- If $L < 1$ the series is convergent.
- If $L > 1$ the series is divergent.
- If $L = 1$ the series may be divergent or convergent.

Now to present the **root test**, suppose we have the series defined by:

$$\sum a_n \tag{12}$$

Then we can define:

$$L = \lim_{n\to\infty} \sqrt[n]{|a_n|} = \lim_{n\to\infty} |a_n|^{1/n} \tag{13}$$

Now the following conditions would hold:

- If $L < 1$ the series is convergent.
- If $L > 1$ the series is divergent.
- If $L = 1$ the series may be divergent or convergent.

0.3 Reviewing Limit Theorems

In this section we shall consider the limiting behaviour of a sequence of random variables given by $\{z_n\}$. We can discuss the various modes of convergence in a step wise manner.

- A sequence of random variables $\{z_n\}$ **converges in probability** to a constant α if for any $\epsilon > 0$ we have:

$$\lim_{n\to\infty} Prob(|z_n - \alpha| > \epsilon) = 0 \tag{14}$$

  Note that the constant α is called the **probability limit** of z_n and can be also written in the following notations as well:

$$plim_{n\to\infty} z_n = \alpha \tag{15}$$

$$z_n \xrightarrow{P} \alpha \tag{16}$$

- A sequence of random scalars $\{z_n\}$ **converges almost surely** to a constant α if we have:

$$Prob(\lim_{n\to\infty} z_n = \alpha) = 1 \tag{17}$$

  This is a stronger condition than the convergence in probability. That is, if a sequence of random variables converges almost surely, then it converges in probability as well.

- A sequence of random variables $\{z_n\}$ **converges in mean square** to α if we have:

$$\lim_{n\to\infty} E[(z_n - \alpha)^2] = 0 \tag{18}$$

- Now in all of the above scenarios our sequence was converging to a constant value, however convergence holds for a target random variable as well. We can say that a sequence of random variables $\{z_n\}$ converges to a random variable z if:

$$\{z_n - z\} \xrightarrow{P} 0 \tag{19}$$

$$z_n \xrightarrow{P} z \tag{20}$$

- Let $\{z_n\}$ be a sequence of random variables and let F_n be the cumulative frequency distribution of z_n. We can say that the sequence $\{z_n\}$ **converges in distribution** to random variable z if the CDF F_n of z_n converges to the CDF F of z at every continuity point of F. This condition can be written as:

$$z_n \xrightarrow{D} z \tag{21}$$

  Note additionally that F is known as the **asymptotic distribution** of z_n.

0.4 Fundamentals of Time Series

In time series analysis we will mostly be dealing with stochastic processes. A **Stochastic process** is basically a sequence of random variables. Now if the index for the random variables is interpreted as representing time, then what we have is essentially a **time series**. Further we note that if $\{z_i\}(i = 1, 2, \cdots)$ is a stochastic process, its **realization** is an assignment for each i of a possible value of z_i. Therefore the realizations of $\{z_i\}$ is essentially a sequence of real numbers. A fundamental point to note about time series is that we only observe the realisations of the stochastic process underlying the time series, only once.
As an example, consider the annual inflation rate of some country between 1950 and 2000. This would essentially be a list of 50 numbers or values which would form **one possible outcome** for the stochastic process underlying the inflation rate variable. If history took a different course, we would have had a different sample of realizations of the same stochastic process. Now if we could observe historical values many times, we could assemble many samples, each containing a different list of 50 inflation rate values. Note that in this case the mean inflation rate for say 1950 would be the mean inflation rate for 1950 across all the historical samples. This kind of a population mean is called the **ensemble mean** and is defined as the average across all possible different states of nature at a given time period.
While it is obviously not possible to observe alternate histories, if we make the assumption that the distribution of the inflation remains unchanged, that is the set of 50 values we observe are all assumed to have come from the same distribution, then we are essentially making a **stationarity assumption**. Further we state **ergodicity** as the level of persistence in the process, that is the extent to which each element will contain some information not available in other elements.

0.4.1 Stationary processes

A stochastic process $\{z_i\}(i = 1, 2, \cdots)$ is said to be strictly stationary if for any given finite integer r and for a set of subscripts: $i_1, i_2, \cdots, i_r$, the joint distribution of $(z_i, z_{i_1}, z_{i_2}, \cdots, z_{i_r})$ depends only on: $(i_1 - i, i_2 - i, \cdots, i_r - i)$ and not on i. What this basically means is that the length of time period lag is what defines the distributional features and not the start or end of the lag. For example, the distribution of (z_1, z_5) is the same as (z_{12}, z_{16}). The distribution of z_i does not depend on the absolute position of i rather on its relative position. We can infer from this statement that the mean, variance and other higher moments remain the same across all i. Now we note some important definitions within this framework.

- A sequence of **independently and identically distributed** random variables is a stationary process that exhibits no serial dependence.
- There are many aggregate time series such as GDP that are **not stationary** because they exhibit **time trends**. Further we note that many time trends can be reduced to stationary processes. A process is called **trend stationary** if it becomes stationary after subtracting from it a linear function of time. Also, if a process is non stationary but its first difference $z_i - z_{i-1}$ is stationary, then the sequence $\{z_i\}$ is called **difference stationary**.

- A stochastic process is said to be **weakly covariance stationary** if:

$$E(\boldsymbol{z_i}) \textit{ does not depend on } i$$
$$Cov(\boldsymbol{z_i}, \boldsymbol{z_{i=j}}) \textit{ depends on the index } j \textit{ and not on } i.$$

The j^{th} order **autocovariance** denoted by Γ_j is defined as:

$$\Gamma_j = Cov(\boldsymbol{z_i}, \boldsymbol{z_{i-j}}) \tag{22}$$

Further we note that Γ_j does not depend on i due to covariance stationarity. Another condition thus satisfied is:

$$\Gamma_j = \Gamma_{-j} \tag{23}$$

We can say the the 0^{th} order autocovariance is nothing but the variance given by:

$$\Gamma_0 = Var(\boldsymbol{z_i}) \tag{24}$$

The corresponding notation for scalar quantities is :

$$\gamma_j = \gamma_{-j} \tag{25}$$

If we take a string of n successive values of the stochastic process $(z_i, z_{i+1}, \cdots, z_{i+n-1})$ then by the rule of covariance stationarity we can say that the $(n \times n)$ covariance matrix is the same as that of $(z_1, z_2, \cdots, z_n)$ and is given by:

$$Var(z_i, z_{i+1}, \cdots, z_{i+n-1}) = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{n-2} \\ \vdots & \cdots & \cdots & \cdots & \vdots \\ \gamma_{n-2} & \cdots & \gamma_1 & \gamma_0 & \gamma_1 \\ \gamma_{n-1} & \cdots & \gamma_2 & \gamma_1 & \gamma_0 \end{bmatrix} \tag{26}$$

Finally, the j^{th} order **autocorrelation coefficient** is given by:

$$\rho_j = \frac{\gamma_j}{\gamma_0} = \frac{Cov(z_i, z_{i-1})}{Var(z_i)} \tag{27}$$

The plot of $\{\rho_j\}$ against the time index j is called a **correlogram**.

- Another important class of weakly stationary processes is the **white noice process** which is a process with zero mean and no serial correlation.

$$\textit{Covariance stationary process } \{z_i\} \textit{ is white noise if}$$
$$E(z_i) = 0 \textit{ and } Cov(z_i, z_{i-j}) = 0$$

Additionally we note that an independently and identically distributed sequence with zero mean and finite variance is called an **independent white noise process**.

References

[1] Pauls Online Notes (Lamar University)

[2] Thomas Finney - Calculus

[3] Fumio Hayashi - Econometrics