Time Series Modelling and Forecasting with Applications in R

Prof. Sudeep Bapat

Shailesh J. Mehta School of Management

Indian Institute of Technology Bombay

Week 09

Lecture 42: Spectral Representation of a Series

Hello all, welcome to this course on time series modeling and forecasting using R. So again, as probably all of you might know, the topic for this week is spectral analysis and Fourier transformations. So, in the last lecture, we gave a bit of an introduction as to what one means by spectral analysis or, let us say, Fourier transformation of a series. Now again, obviously, before I start with any new things, I will again sort of reiterate very quickly what happened in the last lecture. The idea is to transition from a time domain to a frequency domain or a time domain approach to a frequency domain approach.

So, all the models we have seen so far—be it AR, MA, ARMA, or ARIMA—were in the time domain approach. But of course, if you have a practical dataset with you, which is, of course, a time series dataset, and then, let us say, this is again a hypothetical example. So, let us say the series behaves something like that. Then one can clearly see that you have some sort of repetition going on in the series, right? So, you have a certain frequency in that series, okay?

So, why not study the particular series from a frequency domain approach? And this is the whole idea about this week's lectures, where we talk about, let us say, the spectral density of a series, or the other thing is the spectral representation of a series, or Fourier transform, inverse Fourier transform, right? So, all these ideas try to capture or rather try to look at a time series from a frequency domain approach. Now, again, I have sort of drawn this graph here, but then again, just to mention a couple of other points about what exactly one means by a frequency domain approach. It is nothing but, let us say, the series behaves like that.

So, can you sort of superimpose a couple of sine curves and cosine curves on the series to resemble the behavior of the actual practical series itself, okay? So, a bunch of sine plus cosine—of course, some function of sine and cosine—is nothing but the spectral representation or the Fourier representation of the time series, okay. So, in other words, rather than forming a regression of something like yt and its own past values—let us say yt minus 1, yt minus 2, and so on—that we have seen before in, let us say, autoregressive structure or ARMA structure, etc. Now, we will form a certain regression between yt and some sinusoids. So, when I say sinusoids, they mean either some sine functions of, let us say, sine theta or cosine functions, etc.

So, the transition from a time-domain approach to a frequency-domain approach is the essence of this week. Okay, so now moving on. So, the topic for today would be broadly revolving around this idea, which is called a spectral representation and a few points beyond. Okay, so the very first thing we study today is called the spectral representation theorem, or in short, it is SRT. Okay, so again, just to repeat the spectral representation theorem.

So, firstly, we will try to understand what exactly this theorem means, and then there will be again some technicalities here. So, as you see in front of you, you have an equation and so on and so forth. But again, as always, we will try to explain each and every point in a bit more detail. Okay. Now, by the way, just a small disclaimer before I enter into the second lecture of this week is that, again, one very strong suggestion is that the entire area of time series might be new to the majority of you.

So I've been having I've been getting some requests that would this is this course relevant for me. So I come from a certain background. I mean, be it chemical engineering or electrical or civil, etc., So again, so this is almost the end of the course, right? And then still if you're sitting in this course and then putting all your efforts in this course, then obviously it's of very much interest of you.

That being said, of course, time series applications are abundant. So finance is not the only area where time series applications can be found. Of course, you can find time series applications, again, as discussed in the very first lecture in, let's say, climatology, environmental sciences, or even pharmaceutical sciences, ecology, chemical engineering, electrical engineering, civil, etc. So, any data set where you have some time stamps and want to either forecast or try putting forward a model on that data set, you can actually apply some of the elementary analysis or elementary models you discuss in this course. Just a small disclaimer as to should you still go ahead with this course, should you give the examination, should you get the certificate and so on and so forth.

So, all these are clarificatory sort of doubts that people might have. So, coming back to lecture. So, the first idea is the spectral representation theorem. So, firstly we will try to understand as to what exactly this theorem means, right. So, the spectral representation theorem states that any stationary time series can be expressed as a combination of certain sinusoidal functions of different frequencies ok.

And again, each of these sinusoidal functions has its own amplitude and phase. So, all these are technical terms, right? So, frequency, amplitude, phase, right? So, by the way, amplitude stands for the height of each of the oscillations. So, how high is each of the oscillations?

Frequency stands for the length between any two repetitions. So, this is the frequency, let's say, which could be measured in, let's say, pi or 2 pi or in general radians, etc. And then, phase means a certain group of such frequencies. So, let's say if this is the first group, then this is the first phase. So, again, all these are slightly technical terms which are very much relevant in this entire area of spectral representation or Fourier transform, etc.

Anyway, so again, the spectral representation theorem states that any stationary time series, let us say y t, where y t is assumed to be stationary. can be expressed or can be represented as a combination of certain sinusoidal functions of different frequencies. So, if you want to write down the equation, the equation happens to be that. So, y t equals some integral from minus infinity to infinity e to the power i omega t and then z of d omega. Now, again, this d omega is nothing but the integral, and then the integral is with respect to this omega.

$$y_t = \int_{-\infty}^{\infty} e^{i\omega t} Z(d\omega),$$

y_t: stationary time series

ω: angular frequency

$Z(d\omega)$: a complex valued stochastic process which determines the contribution of frequency ω to y_t.

Now, firstly, what is omega? So, omega is the first important parameter, which is nothing but the angular frequency. So, you see omega here in the derivative, as well as here in the exponent. And y(t) is nothing but the stationary time series which we want to express in terms of the spectral representation. And then lastly, this Z(d omega) is a complex-valued stochastic process which determines the contribution of frequency omega to y(t). So, how much this particular frequency omega contributes to that particular instance of the time series y(t) is given by Z(d omega).

So, in short, just to repeat, the equation is y(t) equals the integral from minus infinity to infinity of e to the power i omega t and then Z(d omega). Now, the next important thing we will study is called the spectral density function. So, in the last slide, we simply gave the spectral representation, but now we will define what we mean by the spectral density function. In short, the spectral density function, or S(omega), describes how the variance of a time series is distributed across different frequencies. Again, a couple of important points to note here: the spectral density function S(omega) focuses more on the variance aspect.

So, the variance of the actual time series. So, let us say y(t) is the time series, and rather than focusing entirely on y(t), we are more concerned about how the variance of y(t) behaves or how the frequencies within the variance of y(t) behave. So, the spectral density function describes exactly how the variance of a time series y(t) is distributed across different frequencies. And the second point is this SDF. So, SDF is the spectral density function in short.

So, the SDF provides insight into the periodic components and is widely used let us say in all these areas. So, signal processing, finance and a few other fields as well. Now, again one small disclaimer or one small important point is that wherever you will find applications where you will encounter frequency based time series applications. I mean frequency based means there has to be some repetition, some oscillations like this. So, one can actually transition from a time domain approach to a frequency domain approach and then we can study further about let us say HDF or its Fourier transform, its spectral representation.

So, all these things fall in place ok and these are some of the key areas where one can actually get abundant applications where one can study from a frequency domain approach. So, signal processing. I am not sure how many of you have an electrical engineering background here, but then signal processing means the same thing. So, let us say if you are transporting a signal. So, any signal should have some frequencies like this, right.

So, let us say this is a particular signal and then can we sort of extract the important frequencies in this underlying signal, ok. Or let us say an example from finance. So, let us say if you want to chart out the stock price which is highly volatile. So, let us say this is the returns price or this is the returns scenario. So, can you sort of extract the underlying frequencies in this return scenario etcetera ok.

So, SDF provides insight into the periodic components. So, there should be some periodic components, there should be some periodic oscillations, and hence it is widely used in all these underlying areas, OK. So, now the definition of SDF. So, the spectral density function—how do we write it down and how do we define it, OK? So, for any weakly stationary time series y_t, the spectral density function is nothing but the Fourier transform of the autocovariance function gamma h. So, the last statement is really important: the SDF is nothing but the Fourier transform of what?

$$S(\omega) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) e^{-i\omega h}$$

ω: angular frequency

$\gamma(h)$: auto-covariance function at lag h. $\gamma(h) = E(y_t y_{t+h})$

So, the Fourier transform of the autocovariance function of that time series, and then the autocovariance function is given by gamma h. Now, again, if you are not very comfortable with what you mean by Fourier transform, you might have to revisit the last lecture. So, in the last lecture, we studied Fourier transform, inverse Fourier transform, and how you express or how you write down the Fourier transform—all these things. The only change here is that the SDF is nothing but the Fourier transform of not the time series itself, but the autocovariance function gamma h of the time series. And by the way, if some of you still remember the idea or the definition of the Fourier transform, then this is exactly the definition of the Fourier transform.

So, a general-looking Fourier transform is 1 by 2 pi summation h going from minus infinity to infinity, and whichever entity you want the Fourier transform of, you place it here and then into e to the power minus i omega h. Now, again, a couple of things to note here: this is a notation for a discrete Fourier transform, and why? Because we have a summation here. So, again, within Fourier transform, it could be either discrete or continuous. So, discrete Fourier transform involves summations, while the more general Fourier transform involves integrals. So, I can very well replace the summation by something like integral h going from minus infinity to infinity here.

So, rather than a summation, if I replace it by an integral, it will be a general Fourier transform. If you retain the summation sign, it will be DFT. So, DFT is the discrete Fourier transform. Now, again coming back to the definition: Fourier transform of what? So, Fourier transform of the autocovariance function gamma h. So, here we have this gamma h here into the same thing.

So, e to the power minus i omega h. And again, just to sort of repeat one more time, this term here that you see, right? So, e to the power minus i omega h can be written down as a function of sine and cosine. So, I think we discussed this even in the last lecture, right? Where there is a connection between this term here, this exponent here, and functions of sine and cosine, okay?

And hence, if you want to express any time series into sinusoids, then this function should come here, okay? Now, again, omega is what? So, omega is again the angular frequency; gamma h is the autocovariance function at lag h. So, again by this time, if you might have forgotten how to define gamma h, the definition is gamma h is nothing but the expected value of y t into y t plus h. So, expected value of y t into y t plus h. So, in general, this S omega gives you the spectral density of the autocovariance function of a time series y t. So, now a few properties—some important properties—about this SDF or the spectral density function. So, the first one is, for a real-valued time series, S of omega is always symmetric.

So, what do you mean by symmetric? So, S of omega happens to be exactly equal to S of minus omega. The second point is non-negativity. So, S of omega is obviously non-negative for any and all omega. So, S of omega is greater than or equal to 0 for all omega.

Again, whatever properties you know about density functions—right?—obviously apart from the first one, because a general density function for a random variable need not be symmetric. So, the first point is additional, but let us say non-negativity. So, non-negativity is still true for any density function or PDF, as you say. So, PDF of a random variable X. And since we are again calling this a spectral density function, non-negativity has to be there. And then the last point is total variance.

So, the total variance of the time series is nothing but the integral or sum of the spectral density. So, this is a very famous relationship. So, if you want the total variance of y_t—so, y_t is the actual time series, right?—and here, if you want the variance of the underlying time series y_t, that could be found out from summing up the spectral density. So, summation h equals minus infinity to infinity of gamma h, which is nothing but the integral from minus pi to pi of S of omega, and then d omega. By the way, here this should be S of omega—by the way, this should be S of omega even here and even there.

$$Var(y_t) = \sum_{h=-\infty}^{\infty} \gamma(h) = \int_{-\pi}^{\pi} S(\omega) d\omega$$

So, there is a small typo here. So, instead of gamma h, this would be s omega because, to obtain the total variance of the time series, you have to sum or take the integral of the spectral density. Makes sense so far? So, these are some of the properties of the underlying SDF. Okay.

So, a couple more properties. So, the next one is periodicity. So, what do you mean by periodicity? So, for any discrete time series, S omega is always periodic with a period of 2 pi. Now, firstly, what do you mean by that?

So, S omega repeats itself. So, periodicity means after how much lag or after how much duration it repeats itself, right? So, let us say if you observe a time series which goes like that, So, periodic means after this duration or after this much duration, it repeats itself; it has a tendency of repeating itself, okay. At the same time, the HDF or the spectral density function is periodic with a period of 2 pi.

So, after every 2 pi length, the density function or the HDF would repeat itself, okay? So, the underlying periodicity in S omega is nothing but 2 pi. And last but not least, you also have a certain inverse relationship. So, there exists an inverse relationship between the autocovariance function gamma H and S omega as follows. So, one can actually derive the autocovariance from S omega also.

$$\gamma(h) = \int_{-\pi}^{\pi} S(\omega) e^{i\omega h} \, d\omega$$

So, of course, in the last slide, if you go back a slide or probably one more slide. So, here we are sort of deriving S omega from gamma H, isn't it? And please take note of the exponent here. So, the exponent here is negative, right? So, e to the power minus i omega h. So, what would happen if you try to invert this?

So, if you try to invert this, there has to be certain e to the power i omega h, which is positive, right? So, I can—so if you bring this constant, think of a situation where you are bringing this exponent to the LHS, all right? So, if you bring this exponent to the LHS, there should be e to the power positive i omega h. And this is exactly what you will see. A couple of slides down the line. So, there exists a certain inverse relationship between gamma H and S omega.

So, for some reason, if you want to derive gamma H from the underlying S omega, then this is the relationship. So, gamma H equals the integral from minus pi to pi, and then S omega e to the power i omega H, and then d omega. So, hopefully, the idea about spectral density—the spectral representation of a series—is clear. So, what do you mean by the spectral representation of a series? How do you define the spectral density function?

And what are all the properties of the underlying spectral density function, which is S omega? And lastly, what is the relationship between the underlying spectral density function of a time series and the autocovariance function of the time series Y t? So, what exactly is the interpretation now? So, can you conclude something from the last slide or, in general, from the spectral density? The answer is yes.

So, the first case is of low frequency. So, low frequency means what? So, that omega is really close to 0. If omega is really close to 0, then the frequency is low. Or rather, the angular frequency is low.

So, in such a situation, low frequencies correspond to long term trends or slow moving components in the time series. So, the frequency component omega is low or rather close to 0, then there has to be some really long term trends or slow moving components in the series. So, rather than very fast oscillations, the time series would take its own time to sort of provide a long term trend, something like that. Is the idea clear? So, rather than you will not have something like this where you have more frequent oscillation.

So, this is not the case. So, long term trends means that the time series is really slow moving or you have some slow moving components in the underlying time series. On the other hand for a high frequency, so high frequencies mean omega is close to pi. So, low frequency means omega close to 0 and high frequency means omega is close to pi. And under the situation of high frequencies, so high frequencies represent rapid fluctuations or noise.

So, if the frequency value is almost equal to pi, then one can actually see really rapid fluctuations or more like noise. something like that ok. So, this is when omega is close to pi and let me let me redraw here. So, a situation where you have some slow moving components maybe something like that then this corresponds to omega being close to 0 which is low frequency. Make sense so far?

And then the third interpretation in S omega is that peaks in a spectral density function or S omega indicate dominant periodic components at specific frequencies. Again, just to repeat, so if you tend to find out the peaks, so the peaks in the spectral density function, not the time series. So, by the way, one can actually draw the HDF also. As one can draw any pdf. So as one can draw any general density function of a random variable x similarly I can trace or I can draw or I can curve out the path of the SDF.

So I can draw the SDF also. So, if you tend to draw the SDF, which is S omega, there will be some peaks in that, right. So, those peaks in S omega indicate some dominant periodic components at specific frequencies. So, wherever you have peaks in the graph of S omega, those peaks would indicate dominant components at specific frequency. So, wherever you see significant peaks, I can immediately say that which frequency is dominant, by the way.

So, this is the whole idea. So, by looking at the graph of any spectral density function, one can actually or rather immediately find out the significant frequencies underlying the time series. So, these are some interpretations or some conclusions one can have from the general idea about let us say SDF or S omega etc. So, we will try to elaborate a few examples now. So, the first example is white noise.

So, again, I need not repeat what white noise is. So, white noise is more like—obviously, it is a noise, right? Something like that, where you have irregular fluctuations. So, completely random movement, and this is a particular assumption we are making here. So, let us say y_t is white noise, which is assumed to be normally distributed with mean 0 and variance sigma square. So, again,

One more property of white noise is y_t is iid. So, they have to be independent. So, y_t follows a normal distribution, mean is 0, and then again, variance is fixed at sigma square. So, if you draw a horizontal line here, then this line corresponds to 0, by the way.

So, now again, if you are not very confident about some properties of white noise, you may have to revisit some of the earlier lectures from past weeks, right?

But again, I have a reference here for you: the autocovariance function of a white noise process, which is y_t here, happens to be gamma_h is nothing but exactly equal to sigma square if h is 0, and then autocovariance is 0 for any other value of h. So, let me—if you want to write down gamma_h in a proper manner, it will be something like 0 when h is not 0, or rather, sigma square when h is 0. So, gamma_h is nothing but equal to sigma square when h is 0, and then 0 otherwise. For all other values of h, autocovariance is 0, and whenever h is 0, the autocovariance is sigma square. So, using this idea now, or using this small detail here, can we write down the spectral density?

Of course, yes. So, again, let me go back to the definition or the formula for writing down S of omega. So, S of omega happens to be this here, right? So, again, if you notice this for a second. So, S of omega is 1 by 2 pi summation h is from minus infinity to infinity gamma h into e to the power minus i omega h, ok.

So, let me just write down what would happen for a white noise right here so that you have the formula in front of you. So, again, as seen before, for a white noise process, my gamma H happens to be sigma square whenever H is 0. So, what would be the corresponding S of omega? So, S of omega would be nothing but this constant would still remain. So, 1 by 2 pi and then summation for all possible values of h from minus infinity to infinity there will be 0 there also.

$$y_t \sim N(0, \sigma^2).$$

Auto-covariance: $\gamma(h) = \sigma^2,\ h = 0$ and 0 otherwise.

Spectral density: $S(\omega) = \frac{\sigma^2}{2\pi}$. Which is constant for all ω

So, whenever h is 0, the S of omega would be 1 by 2 pi gamma h. Now, gamma h is sigma square when h is 0 and e to the power minus i omega into 0 which is 1. So, in short, the spectral density whenever H is 0 happens to be that for all the other values the spectral density happens to be 0 because this gamma H here is 0. Make sense? So, it requires a bit of work. I mean, you have to look at the formula of S of omega and try to find out what would happen to S of omega for different values of H. So, again, just to repeat one last time, whenever h is 0, this gamma h is nothing but exactly equal to sigma square.

Whenever h is not 0, gamma h is 0. And then I can immediately write down this S of omega here. So, moving ahead, this is exactly what we have here. So, for a white noise, the autocovariance structure happens to be that. And hence, my spectral density S of omega becomes sigma square by 2 pi.

And now, obviously, there is one very interesting property about this S of omega: that this S of omega is constant for all omega. So, you do not have any variable here. So, sigma square by 2 pi is a constant. What this means, in turn, is that this reflects that the white noise has equal power across all frequencies. Because if you chart out the density, right, if you chart out the S of omega, my S of omega would be purely a constant at this value.

So, sigma square by 2 pi. It will be a horizontal line at that value, which sort of concludes that the white noise has equal power across all the frequencies. Make sense? So, the first example was white noise. Then, the second example could be, let us say, an AR1 process.

So, what happens under an AR1 process? So, this is the AR1 process. So, AR1 means the order is 1. So, y_t equals phi multiplied by y_t minus 1 plus epsilon_t, and epsilon_t follows a normal distribution with mean 0. And variance sigma squared.

$$y_t = \phi y_{t-1} + \epsilon_t,\ \ |\phi| < 1,\ \ \epsilon_t \sim N(0, \sigma^2)$$

Auto-covariance: $\gamma(h) = \frac{\sigma^2}{1-\phi^2}\phi^{|h|}$

Spectral density: $S(\omega) = \frac{\sigma^2}{2\pi}\frac{1}{|1-\phi e^{-i\omega}|^2}$

So, epsilon_t is any other random error we have seen earlier. And this is one condition we require for the process to be stationary, if you remember. So, the absolute value of the coefficient has to be less than 1. Now, again, I can immediately write down the autocorrelation function of this AR1, which takes this form. So, sigma squared divided by 1 minus phi squared multiplied by phi to the power of mod h. And using this gamma h, I can derive the spectral density of AR1, which happens to be that.

And I can sort of conclude towards the end that low frequencies dominate for the absolute value of phi, which is very close to 1, indicating a slow-moving behavior. Okay. So, for a particular AR1 process, the movement is really slow because of all these underlying properties. All right. And lastly, we will talk about the MAQ process.

So, for a moving average process of order Q, this is my Yt here. Again, epsilon t is a random error. I can write down the spectral density, which takes this form. So, sigma squared by 2 pi into the absolute value of this entire thing and then whole squared. A couple of properties are that it is limited to frequencies below a certain level determined by Q. So, once you fix some Q, this could be a drawback here that the spectral density is only limited to those many orders.

$$y_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q}, \;\; \epsilon_t \sim N(0, \sigma^2)$$

$$\text{Spectral density: } S(\omega) = \frac{\sigma^2}{2\pi} |1 + \theta_1 e^{-i\omega} + \theta_2 e^{-2i\omega} + \cdots + \theta_q e^{-qi\omega}|^2$$

So, whatever value Q takes, let us say 4, 5, 6, you will have those many variables here in the spectral density. So, it is limited to frequencies below a certain level determined by the value of Q. However, it is smoother than an AR process of similar order since it is based on past shocks. So, these are some underlying properties when it comes to the spectral density of MAQ versus ARP or AR1, etc. And the last example is a random walk.

$$y_t = y_{t-1} + \epsilon_t, \;\; \epsilon_t \sim N(0, \sigma^2)$$

$$\text{Spectral density: } S(\omega) = \frac{\sigma^2}{2\pi} \frac{1}{|1 - e^{-i\omega}|^2}$$

So, this is a random walk equation. So, y t equals y t minus 1 plus epsilon t, and for the random walk, I can actually derive the spectral density to take that form. And what do you mean by that? So, the spectral density diverges as omega tends to 0, indicating a non-stationary process dominated by low frequencies. So, again, all these conclusions can be verified.

So, this is not the time to verify each and every conclusion. I mean, what happens if omega converges to 0? So, if omega converges to 0, the spectral density diverges, indicating that the underlying series, which is a random walk, is non-stationary, and we already know that. So, we have proved multiple times that a random walk is not stationary. This is one more idea to prove.

So, hopefully by this time, you have some solid understanding about the spectral density function, the Fourier transform, inverse Fourier transform, and the spectral representation of any underlying time series, okay? Now, in the next lecture, we will try to elaborate a bit more on a few other properties where one can transition from a time-domain approach to a frequency-domain approach, okay?

Thank you.