Digital Communication Using GNU Radio

Prof Kumar Appaiah

Department of Electrical Engineering

Indian Institute of Technology Bombay

Week-03

Lecture-15

Hello. So, welcome to this lecture on Digital Communication Using GNU Radio. My name is Kumar Appiah and I belong to the Department of Electrical Engineering at IIT Bombay. In this lecture, we are going to look at demodulation and cover some of the preliminaries and gain some of the tools that will be useful when we perform demodulation, that is to get back your symbols from the waveforms. So as an introduction, you can recall that modulation is the conversion of messages into signal waveforms. So, as you are aware, in the case of digital communication, our aim is to communicate a finite set of messages and these messages have to be converted to waveforms and these waveforms as you have seen are designed with bandwidth and power constraints.

The reverse operation is demodulation, wherein from the waveforms you want to infer or conclude as to which message or messages were sent. So, the task of recovering messages from received signals is demodulation. Here, the tricky part is that since we have waveforms and these waveforms can be very complicated and they can actually undergo variations. As an example, in the case of let us say amplitude modulation, let us say that you send one of one, two, three or four levels, but these levels are actually translated to some waveforms.

So, for example, you can have this is one level, this is another level, this is another level. Now, because of noise, actually what happens is that these things get little bit affected and they do not resemble the original waveform. So, the question is when your waveform undergoes modification under channel noise or other distortions, what do we do? Remember that this is a very practical and typical effect that you have to handle as a communication engineer because the environment in which you are going to send your waveforms is definitely going to cause some modifications. The other thing is that whenever you receive your signal at your receiver or your cell phone or your laptop or wherever it is, it is definitely going to be affected by noise. So, the challenge which we

are going to consider in this lecture are to find out what are the optimal methods to recover messages with minimum errors.

So, we are silently introducing the aspect of error also, that is, it could so happen that you send a particular symbol, but because of the variations caused during the transmission or during reception, it could be misinterpreted as some other symbol. See, after all you are sending a waveform, you are receiving a waveform, what if the received waveform is so far apart or so different that you actually conclude that it is a different waveform, you will end up making an error. Minimizing these errors and ensuring that you get a reliable communication is part of what we call the detection problem. The detection as opposed to estimation refers to the fact that we want to conclude which one of the finite messages you have sent. The ingredients that our toolkit for demodulation has are basically random processes because largely we are going to consider noise here and noise is a random process and it has certain spectral properties that will come in handy.

The other is hypothesis testing, wherein you have to find out which one of these particular hypothesis is most likely given the observation, which is actually another form of or another presentation of the detection problem. Now, let us revise some basics of the Gaussian random variable. Before we actually get to why I am jumping straight into the Gaussian random variable, the reason is because Gaussian random variables appear naturally as noise sources and some channel modification effects as well, which we will cover later, primarily due to the central limit theorem. Now, if you may recall, the central limit theorem essentially says that if you have several iid random variables like

$$X_1 \quad , \quad X_2 \quad , \quad X_3$$

and so on. So, if you add these random variables, then the resulting random variable is actually Gaussian irrespective of the prior distribution that your

$$X_1 \quad , \quad X_2 \quad , \quad X_3$$

had.

For example, you can take

$$X_1 \quad , \quad X_2 \quad , \quad X_3$$

as uniform and add 10 independent and identically distributed uniform random variables and if you actually plot the histogram, surprise, surprise, you will get a bell shaped curve. So, whenever you have variables which appear as a combination of or as a sum of several independent identically distributed random variables, it is more or less accurate

to consider them to be Gaussian and this is confirmed by experimental observations also. Now, since we are going to consider noise significantly in this particular lecture, let us also have an idea as to why this thermal noise or circuit noise is actually generally treated as Gaussian. So, if you look at something like a wire or whatever medium that you are going to consider for sending your signal, typically if it is something like a wire where you are sending signals as voltages and currents, the carriers are electrons. So, you have these little electrons, but these electrons are going to be thermally agitated that is because of the effect of the heat and you know the effect of heat on these electrons, there is a random process which happens.

Some electrons are going to essentially jump or you know go faster or slower and get missing and things like that. That is they essentially superpose some variations onto the current or the voltage that you are going to observe at the receiver. So, since these little tiny variations are all treated as independent, it is a reasonable assumption and they all superpose, that is they add up and because of that you can assume that the thermal noise at the circuit level is typically Gaussian. And this is also borne out by experiments that is a simple thing you can do is connect a wire just from ground to the oscilloscope in your lab and we then just make the oscilloscope go to the most you know like the least possible vertical resolution and take those samples and then transfer it to your PC and histogram it, you will get a bell-shaped curve. So, circuit noise is typically Gaussian.

Now, the PDF of a Gaussian random variable with

mean μ and variance σ^2

that is this is the

mean μ and variance σ^2 .

So, this is something you have to be careful about. The variance is σ^2 , standard deviation σ the mean is μ . This is the PDF, it is

$$\frac{1}{\sigma\sqrt{2\pi}} e^{\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)}$$

by the way here x is a real number. So, the notation which we typically use is

$$N(\mu, \sigma^2)$$

indicating that we have a Gaussian random variable that is this random variable is distributed in a normally distributed that is why we use N or Gaussian, mean is μ , variance is σ^2 .

Now, of course, the PDF you would have seen is bell-shaped. The median sorry for a Gaussian it is interesting that the median which is where the 50 percent distribution lies is mu. The mean is also mu. The mode if you do not recall what a mode is the mode is where the probability density function essentially has a peak. This is called a unimodular distribution because it has a single peak and this peak is at μ .

So, the median, mode, mean all for a Gaussian are μ . The other thing is that the variance the

σ^2

controls how fat your Gaussian is that is how much variability there is. If the Gaussian is narrow then many if you start generating samples of this Gaussian that is if the variance is small you start generating many samples of this Gaussian they will be very close to

μ .

But if your

σ

is large that is your variance is large then you will have variables all over the you will have the random variable realizations all over the place and a wide variance signifies that there is a wide variation you can see. So, it is very typical to assume you know to say that you know people typically say that most of the Gaussian is concentrated within

$\mu \pm 3\sigma$.

If you calculate the probabilities that will be close to 99 percent and so on and you can just do those some people say

5σ , 6σ

and so on. But mostly it is concentrated within let us say a 6σ or you know width of interval with centered around μ . So, that is how a Gaussian looks like. Some other basics of course the sum of these are true for random variables with any distribution. Affine transformations naturally preserve Gaussians this is not true for all random variables of course but still for Gaussians it is true.

If X is distributed as Gaussian with mean μ and variance σ^2

$Y = aX + b$

then Y is distributed with mean with Gaussian with mean

$a\mu+b$

with variance

$a^2\sigma^2$.

Now this is not difficult to see because if you look at

$\mathrm{E}(Y)=\mathrm{E}(aX+b)=a\mu+b$.

Now similarly for variance if you look at the

and I hope you recall that any addition of any constant does not affect the variance. So, this is

$\mathrm{var}(aX)$

and multiplication by a constant essentially results in the variance getting squared and this results in

$a^2\,\mathrm{var}(X)=a^2\sigma^2$.

So, be aware that you must be able to quickly do these transformations.

The reason is because often when we analyze problems related to detection you may want to scale you know if you have

or something you may say oh let me multiply everything by 10 so that becomes easier. So, then you have to make sure that all the distributions are scaled accordingly sometimes you may subtract some value so then you have to adjust the mean. So, make sure that these affine transformations and these computations are very clear. Another thing is this conversion to standard form. So, this N(0,1) is called the standard normal distribution or standard Gaussian distribution that is it is a Gaussian random variable with mean 0 and unit variance that is the variance of this random variable is 1.

Now, the key idea is that any Gaussian random variable there is an affine transformation that will convert this to standard normal. What is it? It is essentially

$\frac{X-\mu}{\sigma}$.

So, why does this work? So, you have the result on top. This is actually like if you want to say use this

$$Y = aX + b$$

you essentially have

$$a = 1/\sigma \quad \text{and} \quad b = -\mu/\sigma$$

that is what you have because it is

$$\frac{X-\mu}{\sigma} \quad .$$

So, that is

$$a = 1/\sigma \quad \text{and} \quad b = -\mu/\sigma \quad .$$

So, it is very evident that

$$a\mu + b \quad .$$

So,

$$a\mu + b = \mu/\sigma - \mu/\sigma = 0$$

Similarly, let us also look at the variance. Let me erase this. So, if you have

$$a = 1/\sigma \quad \text{and} \quad b = -\mu/\sigma \quad .$$

Now, if you want to calculate the variance like I mentioned earlier this b does not play any role in the variance. So, we can essentially ignore this

$$a = 1/\sigma$$

and if you just say

$$a^2 \sigma^2 = \sigma^2/\sigma^2 = 1 \quad .$$

Now, this particular conversion to standard form used to be quite popular in the olden times when we did not have means of computing the Gaussians in a very easy way. So, what would happen is that there would be essentially a booklet or table with lots of these Gaussian probabilities encoded in them and you would essentially have those probabilities encoded only for N(0,1). So, suppose you had a mean of 10 and a variance of 25 you would then convert that whatever Gaussian probability you wanted to find out

for a standard normal by transforming those variables and then just looking it up on the table.

So, this used to be popular in the olden times of course today computers can essentially calculate these things very very easily. Another important aspect that we need to be comfortable with is the CCDF that is the complementary cumulative distribution function of the Gaussian which we also for convenience call the so called Q function. This Q function appears very frequently in the context of digital communication because calculation of symbol errors and bit errors often involve computations over Gaussian of the form this that is probability that your value is larger than you know. So, as we will see shortly or may be in the next lecture or so you typically have to demarket your received symbols into decision regions or decision intervals. So, an error happens if your received value falls outside an interval.

Outside an interval is something like saying

$$Z > X$$

that is just you are integrating and in this case I am using the standard normal. Standard normal is often used because you can always normalize all your computations to make it have the noise to be something like N(0,1) that is something you will see shortly. So,

$$P(Z>x)=\int_{x}^{\infty}\frac{e^{-t^2/2}}{\sqrt{2\pi}}dt \quad .$$

this is the complementary cumulative distribution function for a unit variance 0 mean Gaussian random variable is

$$\int_{x}^{\infty}\frac{e^{-t^2/2}}{\sqrt{2\pi}}dt \quad .$$

Naturally you can see that this is the same as the Gaussian PDF except that you have substituted

$$\mu=0 \quad \text{and} \quad \sigma=1$$

and this for a standard normal is the Q function. Remember one thing the Q is always defined only for a standard normal distribution not for anything which has non-zero mean or a variance other than 1.

Now even this Q computation actually is you can see is a you know it is an integral. So, you will have to compute it numerically or you know do a look up a table like I mentioned. In fact if you are interested in looking at the function in MATLAB, Octave

and you know Python and all those you can look for this ERFC function. ERFC is not a precise function in the same function because ERFC I think does not have some scalings like you know the 2 and 2π you so you will have to do some very minor mapping but the ERFC does this computation. So, this has been implemented for you in many computational software so you can use those each software has a different way of implementing it some of them do a computation some of them do a table lookup and interpolate between values and so on.

But if you really want to just do a quick back of the envelope calculation and approximation you can get some reasonably tight bounds for larger x . So, these bounds are very very useful because you know some sometimes when you are as a designer you want to just get rules of thumb and you want to you know just get quick values you can use some approximations to estimate this Q very very easily. For example

$$\left(1-\frac{1}{x^2}\right)\frac{e^{\frac{-x^2}{2}}}{x\sqrt{2\pi}}$$

and is one of the you know like

$$Q(x)$$

is always larger than this and

$$Q(x)$$

is always smaller than

$$\frac{e^{\frac{-x^2}{2}}}{x\sqrt{2\pi}} \quad .$$

You can see that these two are the same and this just slides slightly lower so you can see that when x starts becoming values like 3, 4, 5 and so on then this becomes a very very tight range of values within which you can sandwich Q. So, Q is essentially approximated by using these.

Now it is worthwhile to see how these can be done you do not need anything more than elementary calculus to do these kinds of computations. So, for example I will show you how you can get one of these bounds. So, you start with

$$Q(x)$$

that is the definition is that it is remember it is the complementary cumulative distribution function for a standard normal. So, a standard normal is

$$\frac{e^{\frac{-x^2}{2}}}{\sqrt{2\pi}}$$

I will write t because I am integrating from x so I will write

$$\int_x^{\infty} \frac{e^{\frac{-t^2}{2}}}{\sqrt{2\pi}} dt$$

that is to visualize it is this part.

Oh I should not run there I am so sorry Gaussian cannot go to 0 this is standard normal yeah it is this part that part always have that picture in mind it is the part to the right of x of course you can draw x to the left also I mean I agree but in general you always take x to be a large positive value so it is the right of x . So, this integral you play a trick you multiply and divide by t you multiply and divide by t should not be a problem because you know even if you say oh what happens if my integrating you know my integration limits have this 0 within it no problem still 0 by 0 does not cause any problem because there the limit is 1 so it does not matter. Now if you play this trick you can just write this as

$$\int_x^{\infty} \frac{d(-e^{\frac{t^2}{2}})}{t\sqrt{2\pi}}$$

and you must remember integration by parts that is

$$\int udv = uv - \int vdu \ .$$

So,

$$uv = \frac{-e^{\frac{-t^2}{2}}}{t\sqrt{2\pi}}$$

you write x and ∞ over here that is you substitute those values

$$-vdu$$

in this case v is you know like this you write it as

$$\frac{e^{\frac{-t^2}{2}}}{t^2\sqrt{2\pi}} \quad .$$

The reason is because you have this $1/t$ and this t that is why you get $1/t^2$ you can check this out.

Now if you substitute over here in the case of ∞ there is no problem it is 0 in the case of x it will be

$$\frac{e^{\frac{-x^2}{2}}}{x\sqrt{2\pi}}$$

which gives you this part

$$-\int_x^{\infty} \frac{e^{\frac{-t^2}{2}}}{t^2\sqrt{2\pi}}\,dt \quad .$$

Now you can clearly see that

$$\frac{e^{\frac{-t^2}{2}}}{t^2\sqrt{2\pi}}$$

is a non-negative function therefore for any value of x this integral will be a positive number since you are subtracting a positive number you can bound this by writing

$$\leqslant e^{\frac{-x^2}{2}} \quad .$$

So, this gives you the left hand side bound over here for the right hand side sorry this gives you the right hand side bound rather yes this is the right hand side bound for the left hand side bound it is not very difficult all you need to do is you need to just continue over here you multiply again by t over here and t over here and if you perform this integral you will get something and you will get something plus. So, you can actually write the greater than or equal to sign and you will get that $1/x^2$ very beautifully. So, you once if you keep going by parts again and again you start getting tighter and tighter bounds of course typically we just restrict ourselves to these two bounds that is the most useful of course emphasizing that these are useful when you want to do and calculations you know but with a calculator but typically nowadays computers can evaluate these very very efficiently.

The next thing that we have to focus on are collections of Gaussian random variables. So, often in digital communication we have to deal with collections of random variables

for example suppose that you are getting multiple symbols at once and for some reason you know there is some relationship between these or some noise related correlations are there and all those kinds of things can happen. So, we do not deal with Gaussian random variables or random variables in isolation we also deal with collections of random variables in the most general form whenever you have a collection of random variables we group them into a vector a column vector typically and call such a vector a random vector. Typically of course there is no need for

X_1 , X_2

to have the same distribution but typically it is convenient for you to group together random variables that have the same distribution not just same mean same variance but the same distribution of course there may be correlations and there may be of course there may be minor variations for example X_1 may be Gaussian with mean 1 variance 2 X_2 may be Gaussian with mean 1 variance 3 X_3 may be Gaussian with mean 10 variance 50 and so on those variations can be there but typically we group together Gaussians in a Gaussian random vector. One important definition or a characterizing aspect of random variables is a so called covariance matrix.

This covariance matrix

$\boldsymbol{C}_x = \mathrm{E}[(\boldsymbol{X} - \mathrm{E}[\boldsymbol{X}])(\boldsymbol{X} - \mathrm{E}[\boldsymbol{X}])^H]$.

So, almost like you want to take the outer product. So, $\boldsymbol{X}$ is a column vector

$\boldsymbol{X}^H$

is an row vector this is the outer product. Now this covariance matrix is a very important quantity that tells you a lot about the joint relationship between the entries in the random vector

$\boldsymbol{X}$

that is

$\boldsymbol{C}_x$

is actually an $n \times n$ matrix whose ij th entry is the covariance between

X_i and X_j

and the diagonal entries the ii th entries are the variance of the corresponding X_i 's. So, what does this mean? So, effectively if you look at the off diagonal entries of $\boldsymbol{C}_x$ these are the variances and this particular entry is the covariance between

X_i and X_j

that is the

$\mathrm{E}[(X_i-\mu_i)(X_j-\mu_j)]$.

So, this is the covariance. Covariance of course you would have seen in your past courses, but to give you an intuition covariance measures both the variance of the individual random variables as well as their correlation. So, how does that work? So, if you look at the next one there is a Cauchy-Schwarz inequality that says the entries of this covariance matrix have to satisfy

$|C_x[i,j]| \leq \sqrt{C_x[i,i]\,C_x[j,j]}$.

This is something which Cauchy-Schwarz inequality says there are many ways to prove this, but more importantly if you look at

$$\frac{|C_x[i,j]|}{\sqrt{C_x[i,i]\,C_x[j,j]}}$$

that will be the correlation coefficient it will be a number between -1 and 1 if you remember the mod will be a number between -1 and 1 and that gives you the correlation coefficient between

X_i and X_j .

Now this is important because obtaining correlations tells us that how much we know about one particular random variable if we know information about the other. Now in a very similar fashion this particular Cauchy-Schwarz inequality also places some constraints on the kinds of entries that can go here.

You can't fill in arbitrary entries and just hope that this becomes a covariance matrix. In other words this particular covariance matrix must satisfy the fact that if you fix the diagonals the off diagonal entries must satisfy this particular condition. So in fact there are some other properties these covariance matrices are by definition you know if for real random vectors they are symmetric matrices for complex random variables they are Hermitian symmetric. You can just put a

C_x^H

over here and you will find that if you just use the relationship

$(AB)^H = B^H A^H$

you get the same

C_x

back. So there are many many properties of these we will explore all those as and when we get to them but for now a covariance matrix characterizes both the variance of the random variables in the random vector and it also characterizes the correlation coefficient.

The correlation coefficient is something which is very useful in deciding how the

X_i s and X_j s

are related. So let's look at collections of random vectors. So when we look at $\boldsymbol{X}$ in this collection of Gaussian random vectors let's say $\boldsymbol{X}$ is distributed as normal and we use the notation

$X \sim N(\boldsymbol{m}_x, \boldsymbol{C}_x)$

where

$\boldsymbol{m}_x$

now is a column vector

$\boldsymbol{C}_x$

is a matrix. So let's say $\boldsymbol{X}$ has n entries, $\boldsymbol{m}_x$ has n entries, $\boldsymbol{C}_x$ is an $n \times n$ symmetric or Hermitian matrix. Now once again if we perform an affine transformation let's say

$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{b}$

then

$\boldsymbol{Y}$

is just going to be distributed again as Gaussian with mean

$\boldsymbol{A}\boldsymbol{m}_x + \boldsymbol{b}$

and covariance

$\boldsymbol{A}\boldsymbol{C}_x\boldsymbol{A}^H$.

So I think it is evident at least the mean part should be evident you can just put an expectation over here and you get

$$\boldsymbol{A}\mathrm{E}[\boldsymbol{X}]+\boldsymbol{b}=\boldsymbol{A}\boldsymbol{m}_x+\boldsymbol{b} \quad .$$

The covariance also be very evident but if you really want to just do it if you write

$$\mathrm{E}[(\boldsymbol{Y}-\mathrm{E}[\boldsymbol{Y}])(\boldsymbol{Y}-\mathrm{E}[\boldsymbol{Y}])^H]$$

then if you just substitute for

$$\boldsymbol{Y}=\boldsymbol{A}\boldsymbol{X}+\boldsymbol{b}$$

then over here

$$\boldsymbol{A}\boldsymbol{m}_x+\boldsymbol{b}$$

comes in. So what you will get is you will get

$$\mathrm{E}[\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{m}_x)(\boldsymbol{X}-\boldsymbol{m}_x)^H\boldsymbol{A}^H]$$

and if you bring the expectation inside you will get

$$\boldsymbol{A}\boldsymbol{C}_x\boldsymbol{A}^H \quad .$$

This is something which you can verify it should not be too difficult. So this is the corresponding affine relationship for the case of Gaussian random variables.

For the case of the Gaussian random vectors like this is the analogous relationship that you found in the random variables case for the random vectors. So let us before ending let us look at some special aspects of Gaussian random variables. If you are given the mean and the covariance the distribution is fully specified like that is okay I mean there is nothing you know that is something which is special for Gaussians but it is something which makes sense. So for example if you look at

$$f_x(X_1, X_2, \ldots, X_n)=\frac{1}{\sqrt{(2\pi)^n|\boldsymbol{C}_x|}}\exp\frac{-(\boldsymbol{X}-\boldsymbol{m}_x)\boldsymbol{C}_x^{-1}(\boldsymbol{X}-\boldsymbol{m}_x)^H}{2} \quad .$$

That is this is something which is special. This is again the matrix or vector version of the corresponding Gaussian. This is like instead of $2\sigma^2$ you have a

$$\boldsymbol{C}_x^{-1}$$

over here you have

$$\boldsymbol{X}-\boldsymbol{m}_x$$

on both sides. So this is the Hermitian should be here and not here apologies yeah. So that

$$(\boldsymbol{X}-\boldsymbol{m}_x)\boldsymbol{C}_x^{-1}(\boldsymbol{X}-\boldsymbol{m}_x)^H$$

gives you a number.

So this is the joint distribution of Gaussian random variables. So there are some couple of properties that is independent implies I mean independent is same as uncorrelated uncorrelated implies independence and joint Gaussianness is preserved and refined transformations. These are a couple of properties. So this and some more properties of Gaussian random variables we will cover in the next lecture. So what have we seen so far we have seen that Gaussian random variables have some special properties and especially when you look at Gaussian random vectors then the distribution is specified when you specify the mean and covariance.

In the next lecture we will cover a little more about the relationship between these jointly distributed random Gaussians as well as hypothesis testing. Thank you. you