Lecture 9 Script

9.1 Graphons

Slide 1: Graphons - Title Page

1. In this lecture we introduce graphons to study graph filters and GNNs in the limit of graphs with very large numbers of nodes.

Slide 2: Graphon definition

1. A graphon is a bounded symmetric measurable function W mapping points of the unit square to the unit interval.
2. This definition is such that we can think of graphons as weighted symmetric graphs with an uncountable number of nodes.
3. The labels of the nodes are the arguments of the graphon functions. That is, the values that x can take in the unit interval.
4. The weights of the edges are the graphon values. The value W(x,y) that the graphon function takes. Observe that since the function is symmetric W(x,y) is the same as W(y,x) for node labels x and y.

Slide 3: Graphon examples

1. To gain some intuition we present here three examples: The uniform or Erdos-Renyi graphon. A symmetric stochastic block model graphon. And an asymmetric stochastic block model graphon.
2. The uniform or Erdos-Renyi graphon is such that W(x,y) is constant and equal to p for all values of x and y. This graphon is related to the Erdos Renyi family of random graphs, which consists of graphs where edges are drawn independently with the same probability.

3. A balanced stochastic block model graphon takes values W(x,y) = p when x and y are both between 0 and ½ or both between ½ and 1. This is signified by the dark blue areas in the figure. The SBM graphon takes values W(x,y) = q when either x or y are between 0 and ½ and the other argument, y or x, is between ½ and 1. This is signified by the light blue areas in the figure. The value p is much larger than the value q. This graphon models two communities, one community with labels varying from 0 to ½ and the other community with labels varying from ½ to 1. The connection within communities is strong. The edge weights are large. The connection across communities is weak. The edge weights are small.

4. An unbalanced SBM graphon has an analogous definition except that the sizes of the communities are unbalanced. One community is larger than the other. In the figure we have a community that spans labels from 0 to 0.2 and another community that spans labels from 0.2 to 1. Within communities we have strong weights p. Across communities we have weaker weights q.

Slide 4: The Purpose of a Graphon

1. What is the purpose of defining these graphs with uncountable nodes? Philosophically, phenomena are easier in uncountable spaces. This is the reason humanity invented calculus. But what is the phenomenon we are trying to simplify?

2. In practice, graphons are abstractions for families of graphs with large numbers of nodes in which members of the family have similar structure. Instead of studying individual members of the family, we study the graphon. This is likely easier. And it also provides information about the whole family. Not individual graphs.

3. In theory, the graphon provides a generative model for graph families. We can generate graphs through stochastic or deterministic sampling of the graphon. These sampled graphs share a common property. Which is that they are sampled from the same graphon.

4. Another, more subtle, epistemological value, is that graphons are limit objects for sequences of graphs. As the number of nodes of the graph increases, it is intuitive to expect graphs to approach a limit. The graphon is this limit. Let us dwell deeper into these three comments,

Slide 5: The Product Similarity Graphon

1. In terms of the practical value of graphons, recall our encounters with product similarly graphs. As we look at these graphs for different collections of products, we see that these graphs “look like each other.” This is true even if the number of products wea re considering is different.

2. We can therefore abstract these similarities into a limit object. This limit object is what we would call the “product similarity graphon.” The reason for creating a limit abstraction is that similarities are more apparent for larger graphs. We can think of product similarity graphs as elements of a sequence converging to a product similarity graphon that encodes this shared structure.

3. It is important to point out that this limit graphon is **not** something that we compute in practice.

4. We just use the abstract idea of a graphon to work with all of these graphs as if they were the same object. Which in a sense they are. They are all close to the “product similarity graphon.”

Slide 6: Graphons as Generative Models

1. To use graphons as generative models it suffices for us to consider samplings of the [0,1] interval.

2. To generate the vertices of a graph with n nodes we sample n points u_1 through u_n from the unit interval.

3. These points can be sampled in a number of ways. Most often we take them from a regular partition of the unit interval or we sample uniformly at random from the unit interval.

4. In either case, each sample corresponds to a node label of the graph. We can use u_i itself as the label. Or we can use i as the label. Or any other one for that matter. We know that the names of labels are not important.

5. What is more important in the generation of the graph is the determination of edge weights. For doing that we evaluate the weight W(u_i, u_j) to determine the properties of the edge i-j. There are two ways in which we can use W(u_i, u_j) to generate graph edges.

6. In the stochastic model we create an edge connecting i and j with probability W(u_i, u_j). If we add this edge, the edge is unweighted and undirected.

7. In the weighted model we add an edge connecting i and j whenever the graphon function W(u_i, u_j) is not null. When we add this edge the edge is undirected but we now make it weighted. The weight of the edge is the graphon function value W(u_i,u_j).

Slide 7: Uniform Graphons as Generative Models

1. As an example of the use of graphons as generative models, consider the uniform graphon.

2. We can use the uniform graphon.

3. To generate uniform random graphs with the same

4. Or with a different number of nodes. The specific edge structure of the generated graphs differs across instantiations. But they have a shared structure that they receive from the generating graphon.

Slide 8: Balanced SBM Graphons as Generative Models

1. This latter observation is more clear if we consider further examples. For instance, consider the balanced SBM graphon.

2. As we did with the uniform graphon, we can use this balanced SBM graphon

3. To generate balanced SBM random graphs with the same.

4. Or with a different number of nodes. Specific edge structures differ across instantiations. But this graphs have a common structure that is caerly different from the structure of the uniform random graphs we have just seen.

Slide 9: Unbalanced Stochastic Block Model (SBM) Graphon, Stochastic Graph Samples

1. Likewise, if we consider the unbalanced SBM graphon.

2. We generate unbalanced SBM random graphs with the same.

3. Or with a different number of nodes. These graphs are different from each other. But they are more similar to each other than they are to the graphs we generated with the balanced SBM graphon or to those we generated with the uniform graphon. They belong to a family we have generated using the unbalanced SBM graphon.

Slide 10: Uniform Graphons as a Limit Object

1. The other, more subtle value of graphons, is their use as limit objects of sequences of graphs. In the figure we consider random graphs with increasing number of nodes. It is clear that as the number of nodes grows, the graph is approaching some sort of limit.
2. But it is unclear what that limit is.
3. We will see in this lecture, that the graphon is the limit. That the graphon is introduced to formalize the notion of convergent graph sequences.

9.2 Convergence of Graph Sequences

Slide 11: Convergence of Graph Sequences - Title Page

1. We consider graph sequences and introduce the notion of convergence in terms of homomorphism densities. We use this convergence notion to define graphons as the limit objects of convergent sequences of graphs.

Slide 12: Convergent Graph Sequences

1. Consider a sequence of graphs G_n with growing number of nodes n. Each of the graphs in the sequence is characterized by a set of vertices V_n, a set of edges E_n, and an adjacency matrix S_n. The graphs in the sequence may be weighted or not. Thus, the entries of S_n need not be binary. They can be general edge weights. We also point out that we are using S_n, harking back to the concept of graph shift operator. But we are restricting our attention to adjacency matrices. S_n is not an arbitrary matrix representation of the graph.

2. We have already intimated that we want to study graph sequence G_n that converges to a graphon W as n goes to infinity. This is illustrated by the sequence of uniform random graphs on the bottom of the slide, which converge to the uniform graphon. But, and as we already intimated as well, it is unclear how this sequence converges and in what sense it converges to the graphon.

3. To understand convergence of graph sequences to graphons, we have to begin with the introduction of three concepts: Motifs, homomorphisms, and homomorphism densities.

Slide 13: Motif and Graph Homomorphisms

1. A motif F is a graph. It can be any graph, but it is convenient to think of it as a small graph that we could embed into another, larger graph.

2. An example of a motif is this star graph we show on the left. Made up of a center node and three satellite nodes. This motif can be embedded into the larger graph as we show in the right. It is important to observe that the motif can be embedded into the graph in multiple ways.

3. This is another place where the star motif can be embedded

4. This is another.

5. This is another. There are many other more ways in which we can embed the star motif into this graph.

6. Another possible motif is this cycle graph with four nodes. This motif can be embedded into the large graph on the right as shown.

7. It can also be embedded in this other place

8. Or this other place

9. Or this other. As was the case of the star graph in can be embedded in multiple places. An important observation is that the number of ways in which we can embed the cycle motif, is, in all likelihood, different from the number of ways in which we can embed the star motif

10. Another example of a motif is this hexagon motif. Which is in fact a cycle with 6 nodes. The hexagon motif can be embedded as shown

11. And it can also be embedded here

12. And here

13. And here. Likely, the number of ways in which we can embed the hexagon motif is different from the number of ways in which we can embed the cycle or star motif.

14. Before we move onwards we need to define homomorphisms formally. A homomorphism is an adjacency preserving map from a motif F into a graph G. Thus, If the motif has vertices V-prime and edges E-prime while the graph has vertices V and edges E.

15. A homomorphism is a map beta from the **nodes** of motif F into the **nodes** of graph G.

16. Such that if i-j is an **edge** in the motif F, an element of the edge set E-prime.

17. Then, the map images beta-of-i and beta-of-j are an edge of the graph G. An element to the edge set E. It is ready to verify that this is true of the homomorphisms we have illustrated.

Slide 14: Homomorphism Count

1. As we have already emphasized, a motif F can be embedded into a graph G in multiple ways. There are multiple homomorphisms from a motif F to a graph G.

2. For the star graph, this is a possible homomorphism function.

3. This is a second possible homomorphism.

4. This is a third.

5. And this is fourth. There are may more.

6. The count of all homomorphism functions is the quantity hom-of-F-G. It is the total number of ways in which we can embed the motif into the graph.

Slide 15: Homomorphism Density

1. Related to the notion of homomorphism count, is the notion of homomorphism density. To define this, observe that If the graph G has n nodes and the motif F has n prime

nodes, there are a total n to the n prime different maps from F to G. Only a fraction of these total number of maps are homomorphisms.

2. We therefore define the homomorphism density of the motif F into the graph G as the fraction of maps that are homomorphisms.

3. That is, the ratio between the homomorphism count hom-of-F-G and the total number of maps n to the n-prime.

4. We denote the homomorphism density as t(F,G). This is a **relative** measure of the number of ways in which the motif F can be mapped into the graph G while preserving the adjacency structure of the motif. It is therefore a quantity that we can expect to settle into a limit. As the size of the graph grows, the number of possible homomorphisms grows. But so does the total number of possible maps.

Slide 16: Homomorphism Density for Weighted Graphs

1. The definition we have just given can be extended to weighted graphs. The extension is not complicated. But it gets a little cumbersome. Consider a graph with adjacency matrix S. Whose entries are not necessarily binary.

2. The homomorphism density of motif F into the weighted graph G is defined as shown. This definition still divides by the total number of possible maps, n to the n-prime. But the numerator is a different expression.

3. The expression looks quite different from the homomorphism count but it is actually quite similar. We are still counting the total number of homomorphisms. This is what the sum over different beta signifies. But each homomorphism is weighted by the product of the edge weights in the homomorphism image. If we map the motif into large edge weights, we multiply by larger numbers. We add a large value to this weighted homomorphism count. If me map the motif into small edge waitsm we multiply by smaller numbers. We add a small value to this weighted homomorphism count.

Slide 17: Homomorphism Density for Graphon

1. We still have a third definition of homomorphism density to introduce. This is the homomorphism density of a graphon.

2. This is akin to the definition that we have just introduced for weighted graphs except that the sum is replaced by an integral. Observe how in the definition the products are over the edges and vertices that define the motif.

3. This integral has a very simple interpretation. It is the probability of drawing the motif F from the graphon W when we sample n-prime nodes from the graphon.

Slide 18: Convergence in Homomorphism Density Sense

1. With these definitions in place we can now define convergent graph sequences.

2. A sequence of undirected graphs G_n converges to the graphon W if and only if for all motifs F

3. The homomorphism density of motif F into the graph G_n converges to the homomorphism density of motif F into the graphon W as n goes to infinity.

4. Convergence of graph sequences to a graphon entails convergence of the homomorphism densities of all motifs. We therefore say that the sequence G_n converges to W in the sense of homomorphism density.

5. An important aspect to emphasize of this definition is that every graphon is the limit object of a sequence of convergent graphs

6. The complementary observation, namely, that every convergent graph sequence converges to a graphon is also true.

Slide19: Example of Convergent Graph Sequence

1. For an example of convergent graph sequences, consider a sequence of random graphs G_n drawn from the graphon W as shown on the bottom of the slide.

2. Each graph G_n has labels u_i drawn uniformly at random from the unit interval

3. And the edge set is such that nodes u_i and u_j are connected with probability W(u_i,u_j).

4. It can be shown that this graph sequence converges to the graphon in the homomorphism density sense with probability 1. The figure on the slide illustrates a

stochastic graph sequence drawn from the uniform graphon with a growing number of nodes. The limit of this sequence is the uniform graphon from which the graphs are drawn. The almost sure convergence of the random graph sequence to the graphon from which the graphs are drawn holds not only for this particular choice of graphon. It holds for all graphons.

Slide 20: Induced Graphons

1. We close this video with a concept that will be useful in later discussions. This is the notion of the graphon induced by a graph. The point is that every undirected graph admits a graphon representation, which we call its induced graphon.

2. Formally, consider a graph G with n nodes and a graph shift operator S in which the weights have been normalized to be between 0 and 1.

3. We construct a partition I_1 through I_n of the unit interval. In which the intervals I_i are regularly-spaced. The i-th interval goes from (i-1) over n to i over n. This is what is called a regular partition of the unit interval with n subintervals. Observe that the partition subintervals are closed on the left and open on the right. Except for the last partition that is closed on the left and the right. We do not write this latter fact to avoid complicating an otherwise simple definition.

4. We define the induced graphon W_G by assigning weight S_ij to the image of the graphon on the Cartesian product between intervals I_i and I_j. That is, if the argument u belongs to partition I-sub-i and the argument v belongs to partition I-sub-j, we assign the value S-sub-i-j to the induced graphon. The weight that corresponds to the edge that matches the cardinality of the intervals. The figure illustrates this construction for the graphon induced by the cycle graph with 6 nodes. The colored regions are the parts of the graphon to which we assign nonzero values. Each of them, corresponds to one of the edges of the cycle.

9.3 Graphon Signals

Slide 21: Graphon Signals - Title Page

1. In this section we introduce the concept of graphon signals. Graphon sigals are signals supported on graphons and their value lies in that they are limit objects of graph signals.

Slide 22: Graphon Signals

1. A graphon signal is a pair (W, X) in which W is a graphon and X is a function mapping the unit interval to the real numbers.

2. The function X in the graphon signal is required to have finite energy. This is equivalent to saying that X belongs to the space of L_2 functions supported in the unit interval.

3. In case you are not familiar with the meaning of finite energy for continuous signals, it just means that the integral of the square of X is bounded. The figure on this slide shows an example of a graphon signal. It involves a function supported in the unit interval, and a graphon. As is the case of graphons, graphon signals have dual interpretations:

4. They can be seen as generative models of graph signals.

5. And they can be seen as limit objects of convergent sequences of graph signals

Slide 23: Graphon signals as generating models

1. The use of a graphon signal as a generative model involves the generation of graph signals (S_n ,x_n) by taking n samples of the graphon signal (W,X).

2. This sampling process involves the selection of labels u_i and the sampling of the graphon at these labels. The sampling can be stochastic or weighted. Same as when we sample graphs from graphons

3. The difference is that now we add a samping of the function X at node labels u_i. This generates the values of the graph signal associated with corresponding nodes.

4. An important point to emphasize in this definition is that the sampling of the graphon and the function X must be at the same labels. The sampling must be consistent.

Slide 24: Induced Graphon Signals

1. To explain graphon signals as limit objects we need the notion of induced graphon signals. This is a definition that we obtain by leveraging the definition of induced graphons.

2. Formally, every graph signal x supported on a graph G induces a graphon signal X_G supported on the induced graphon W_G.

3. To obtain the induced graphon signal we consider the regular partition of the unit interval with n nodes. The same partition we used for the induced graphon. We have n subintervals that are of equal width 1-over-n.

4. We obtain the signal component of the induced graphon signal by making X_G of u equal to x_i for all the arguments u that lie in the i-th interval of this regular partition. As we illustrate in the figure on the right we extend the graph signal value to cover the whole of the i-th partition. Recall that the subintervals of the partition are closed on the left and open on the right. Except for the last partition that is closed on the left and the right.

5. The graphon components of the induced graphon signal is the graphon induced by the graph G, which we have already defined.

6. For completeness, we recall that this induced graphon W_G assign the value S-sub-i-j when the argument u belongs to partition I-sub-i and the argument v belongs to partition I-sub-j

Slide 25: Convergent sequences of graph signals

1. We can now introduce graphon signals as limit objects of graph signals.

2. A sequence of graph signals G_n comma x_n is said to converge to the graphon signal (W,X)

3. If there exists a sequence of permutations pi_n such that for all motifs F we have that:

4. The graph sequences converges to the graphon in the sense of homomorphism densities. Namely, the homomorphism density of the motif F into the graphs G_n converges to the homomorphism density of motif F into the graphon W for all motifs F. This is just convergence of the graph sequence to the graphon.

5. The novel aspect of the definition is to add convergence towards 0 of the L_2 norm of the difference between the graphon signal X and the signal X_sub pi_n of G_n. Namely, the graphon signal induced by the graph G_n relabeled according to the permutation pi_n.

6. We say that the graphon signal (W,X) is the limit of the graph signal sequence (G_n, x_n).

7. The permutation is used in this definition to make convergence independent of labels. This is not needed in the definition of convergence of graph sequences because homomorphism densities are independent of labeling. To retain label independence when comparing signals, we need to incorporate the proper permutations that make the signals as close as possible. This is the same familiar notion of distances modulo permutation which we encountered in our stability analysis of graph filters and GNNs.

8. Further note that our goal is to compare the **vector** x_n with the **function** X. This is an apples to oranges comparison that we resolve with the use of the induced graphon signal. We do not compare the **vector** x_n with X. Rather, we compare the function X_sub pi_n of G_n induced by x_n. This is an apples to apples comparison. We compare two functions.

Slide 26: Graphon shift operator

1. The fundamental operation that we perform on graph signals is a multiplication with the shift operator. The analogous of that in graph signals is the application of the integral linear operator T-sub-W associated with with the graphon W. This linear operator is a functional that maps graphon signals to graphon signals.

2. When applied to the graphon signal X, the operator T_W

3. Produces the signal T_wX whose value at v.

4. Equals the integral from 0 to 1

5. Of the product between the graphon W of u-v

6. And the graphon signal X-of-u.

7. T_W is what we call a Hilbert-Schmidt operator. This is because W is bounded and compact.

8. THe most important point to remark is that this operation is conceptually the same as a matrix multiplication. It is, in fact, the limit of a matrix multiplication. Integrals are not sums. But it is always helpful to think of them as sums.

9. Drawing on the proximity of the operator T_W with matrix multiplications, we call it here the graphon shift operator of the graphon W.

10. Applying the WSO T_W to the graphon signal X has the effect of diffusing the signal X over the graphon W. We know that this is true intuitively because the definition of the WSO is akin to a matrix multiplication. We will show that it is true formally when we get to studying graphon filters.

9.4 The Graphon Fourier Transform

Slide 27: The Graphon Fourier Transform - Title Page

1. Our next goal is to generalize notions and concepts of graph signal processing to graphons. We start with the definition of a Fourier transform for graphon signals.

Slide 28: Eigenfunctions and Eigenvalues of the Graphon Shift Operator

1. By definition, a graphon W is a bounded and symmetric measurable function. That makes it possible for us to associate with it the graphon shift operator T_w, which is a self-adjoint Hilbert-Schmidt operator.

2. This operator is such that when applied to the signal X, it produces the signal T_w-X whose value at v is:

3. The integral of the product between the kernel of the operator and the function X. From this definition we conclude that the operator T_W is bounded. This is because the kernel —- the graphon W —- is bounded. We also conclude that the operator T_W is self adjoint. Because the graphon W is symmetric.

4. When given operators of this type we can define eigenvalues and eigenfunctions. We say the function phi from the interval [0,1] to R is an eigenfunction of T_W with associated eigenvalue lambda if

5. After applying the graphon shift operator T_w to the function phi

6. We retrieve a scaling of the original function. Thus, as the name suggests, an eigenfunction is a generalization of the concept of eigenvectors to a functional space.

7. A graphon shift operator Tw has an infinite but countable number of eigenvalue-eigenfunction pairs lambda i, phi i. This is to be contrasted with the fact that the graphon trakes values on the interval 0-1. Which is uncountable. This property, namely, that the eigenvalues are countable, will prove very important in our theory.

8. For future reference we observe that we assume the eigenfunctions to be normalized to unit energy. That is, that the L2 norm of an eigenfunction phi-sub_i is equal to 1.

Slide 29: Eigenfunctions and Eigenvalues of the Graphon Shift Operator

1. The eigenfunctions phi i of the graphon shift operator Tw form an orthonormal basis of the space L2 [0,1]. That property will be fundamental to construct a graphon Fourier transform capable of decomposing a graphon signal on the basis made up by the eigenfunctions of the graphon shift operator, as we will see in a few minutes.

2. At this point it is important to observe that since the kernel of a self-adjoint integral Hilbert-Schmidt operator can be decomposed in the operator basis, we can thus decompose the graphon W in the basis of eigenfunctions of the operator Tw. In particular, we can rewrite W as a sum over the product of an eigenvalue lambda i and associated eigenfunction phi_i.

3. That is similar to the eigenvector decomposition of a graph shift operator S, which, as we have seen earlier in the course, can be decomposed as the product between a matrix V made up by the eigenvectors of S, and a diagonal matrix Lambda containing the eigenvalues of the graph shift operator.

Slide 30: The Range of the Graphon Eigenvalues

1. As we said, the graphon shift operator Tw is self adjoint, symmetric, and defined on the unit interval. That implies that the eigenvalues of a graphon are real and lie on the interval minus 1, 1.

2. We choose to order eigenvalues with negative indices j in decreasing order, and positive indexes j in increasing order. As we show in the figure, the eigenvalue lambda_1 is the largest positive eigenvalue. As we increase the index, we move towards smaller positive eigenvalues. The eigenvalue lambda_minus-1 is the largest negative eigenvalue. As we decrease the index, as we increase its absolute value, we move towards smaller negative eigenvalues. All positive eigenvalues have a positive index and all negative eigenvalues have a negative index.

Slide 31: Eigenvalues Concentrate Around Zero

1. The most important point for us to observe abut the eigenvalues of a graphon is that thye accumulate at the point lambda equals 0. That is, the eigenvalues converge to zero as j tends to plus or minus infinity.
2. And This is the only point of accumulation for eigenvalues.
3. A consequence of this fact is that for any constant c that we fix, the number of eigenvalues lambda j that have absolute value larger than c is finite.
4. Another important property to point out is that all eigenvalues that are **not** 0 have finite multiplicity.

Slide 32: Eigenvalues of a Convergent Graph Sequence Converge to Those of the Graphon

1. As we can see graphons as the limit objects of convergent graph sequences, it is not unreasonable to expect the eigenvalues of a convergent graph sequence to converge to the eigenvalues of the limit graphon. This result is formalized in the following theorem.
2. According to the theorem, if a sequence of graphs Gn
3. Converges to a graphon W
4. In the homomorphism density sense
5. Then
6. (empty)

7. If we take the limit as n tends to infinity
8. Of the ratio between an eigenvalue lambda j of Sn, and n
9. That is equal to the to the eigenvalue lambda j of the limit graphon W
10. Which corresponds to the limit as n tends to infinity of the eigenvalue lambda j of the graphon induced by Gn
11. Which holds for all j
12. That is, the theorem states that, for any convergent graph sequence Gn, the eigenvalues lambda j of the graph shift operator Sn converge to the eigenvalues lambda j of Tw, with Tw the graphon shift operator associated to the limit graphon W

Slide 33: Eigenvalues of a Convergent Graph Sequence Converge to those of the Graphon

1. According to the previous theorem, for any convergent graph sequence
2. The eigenvalues of the graph, in blue
3. Converge to those of the limit graphon, in red
4. Note that, as expected from our earlier analysis, the eigenvalues of the graphon accumulate around zero —- which does not hold for the eigenvalues of the graph.
5. More precisely, convergence of the eigenvalues holds in the sense that
6. There exists some n 0
7. Such that
8. For every n larger than n 0
9. The absolute value
10. Of the difference between the eigenvalue of the graph shift operator of Gn
11. And the eigenvalue of the limit graphon

12. Is less than or equal to some epsilon
13. But the value of n_0 for which convergence holds is different for each j
14. Thus, the convergence of the eigenvalues is not uniform

Slide 34: The Graphon Shift Operator Induces a Transform [if keeping this slide]

1. The decomposition of the graphon on the operator's basis allows us to rewrite the graphon shift operator Tw as
2. A sum over the product between an eigenvalue lambda j, the associated eigenfunction phi j and the integral of the product of that eigenfunction phi j and the original graphon signal X
3. The integral terms in that expression correspond to inner products <X, phi j> between the signal and a particular eigenfunction, which we can see as a projection of the original graphon signal X over that particular eigenfunction
4. But we saw that the eigenfunctions of a graphon shift operator form a complete orthonormal basis of L2 ([0,1])
5. Thus, those inner products can provide a complete representation of the original graphon signal X on the basis of the graphon shift operator
6. That change of basis without loss of information about the signal is what we call the graphon Fourier transform. Similar to how a graph Fourier transform decomposes a graph signal into the eigenfrequencies of the graph shift operator.

Slide 35: The Graphon Fourier Transform (WFT)

1. Given that motivation for the definition of a graphon Fourier transform, we now proceed to define that concept formally
2. The graphon Fourier transform (WFT)
3. Of a graphon signal X
4. Can then be defined another graphon signal X hat defined over the operator's basis

5. Where each component X hat j of that signal

6. Corresponds to the component of the signal associated to a particular frequency defined by the eigenvalue lambda j

7. And can be computed as the integral from 0 to 1

8. Of the product between the original graphon signal X

9. And the corresponding eigenfunction phi j

10. With lambda j the eigenvalues and phi j the eigenfunctions of the associated graphon shift operator Tw

11. Since the eigenvalues — and eigenfunctions —- are countable

12. The graphon Fourier transform X hat can always be defined

Slide 36: The Inverse Graphon Fourier Transform

1. Naturally, we next define an inverse graphon fourier transform that maps signals defined on the graphon shift operator's basis back to the original domain

2. The inverse graphon Fourier transform (iWFT)

3. Of a graphon signal X hat

4. Can then be defined as

5. The sum over the countable indexes j

6. Of the product between the component of X hat

7. Corresponding to the eigenvalue lambda j

8. And the associated eigenfunction phi j

9. With lambda j the eigenvalues and phi j the eigenfunctions of the associated graphon shift operator Tw

10. Since the eigenfunctions phi j form a complete orthonormal basis of L2 ([0,1])

11. We can see that the inverse graphon Fourier transform retrieves the original graphon signal X without loss of information. Hence, the iWFT is a proper inverse of the graphon Fourier transform.

9.5 The GFT Converges to the WFT

Slide 37: The GFT converges to the WFT

1. In this part of the lecture we discuss the convergence of the **graph** Fourier transform to the **graphon** Fourier transform for graphs sequences that converge to graphons.

2. Doing so requires that we review convergence results for sequences of graphs that converge to graphons.

Slide 38: The Graphon Fourier Transform and the Graph Fourier Transform

1. The graphon Fourier transform of a graphon signal W-X is a projection of the signal X in the eigenspace of the graphon W. Component j of the WFT is the inner product of the signal X with the eigenfunction phi_j. This is the j-th eigenfunction of the graphon, which is associated with eigenvalue lambda_j.

2. The **graph** Fourier transform of a **graph** signal G_n-x_n is a projection of the signal x in the eigenspace of the graph G_n. Component j of the **G**FT is the inner product of the signal x with the eigenvector v_n-j. The eigenvector of G_n that is associated with eigenvalue lambda_j

3. Given the similarity of these two definitions, it is reasonable to conjecture that if we have a sequence of graph signals G_n-x_n that converges to the graphon signal W-X the corresponding sequence of **graph** FTs converges to the **graphon** FT.

4. This conjecture gets more credence if we remember that eigenvalue convergence holds. The eigenvalues lambda_n-j approach the eigenvalue lambda_j.

5. This conjecture should hold

6. Alas, the conjecture is wrong.

7. Convergence of the GFT to the WFT does **not** hold in general. It is not true that the sequence of GFTs converges to the WFT.

8. The reason for our conjecture to fail is that for the GFT sequence to converge to the WFT, we need to state convergence of eigenvectors to eigenfunctions.

9. Convergence of eigenvalue sequence is not sufficient to claim convergence of the GFT sequence because the GFT and the WFT are projections on eigenvectors and eigenfunctions. They are not projections on eigenvalues.

Slide 39: Convergence to Graphon Eigenvectors

1. The challenge in claiming convergence to graphon eigenvectors is that convergence is affected by how close the eigenvalues of other eigenvectors are.

2. Suppose that we consider convergence towards the eigenvector associated with lambda_2 in the figure. We know that the sequence of lambda_2 eigenvalues associated with the sequence of graphs converges to lambda_2. We expect the eigenvector sequence to converge as well. This is true. The eigenvector converges.

3. But how deep we need to go into the sequence index n to observe that we approach the limit, depends on how far lambda_2 is to other eigenvalues. In this particular case, how far lambda_2 is from lambda_1 and lambda_3. This is not a problem for the convergence of the eigenvector associated with lambda_2. Eigenvalues lambda_1 and lambda_3 are not close to lambda_2.

4. But it is a problem for eigenvalues lambda that are close to zero. The eigenvalues of the graphon accumulate at zero. All eigenvalues in the sequence of graphs converge to some graphon eigenvalue. But other eigenvalues of the graphon are close.

5. This makes the eigenvectors slow to converge. They all converge, but convergence is not uniform.

Slide 40: Eigenvalue Margin for Linear Operators

1. To state eigenvector convergence formally we introduce the notion of eigenvalue margin for linear operators. Consider eigenvalues lambda_j of the graphon W and lambda_n-j of the graph G_n. This graph is part of a sequence that converges to W, but we don't need

that information in this definition. It is important to notice that both eigenvalues have the same index, though.

2. We then proceed to compare the **graphon** eigenvalue lambda_j with the closest **graph** eigenvalue lambda_n-i other than lambda_n-j.

3. As we show in the figure, we grow an interval around the **graphon** eigenvalue lambda_j. The interval grows until it hits the first **graph** eigenvalue lambda_n-**i** that is not lambda_n-**j**. We call this margin d_1.

4. We do the same for the graph. That is, we compare the **graph** eigenvalue lambda_n-j with the closest **graphon** eigenvalue lambda-i other than lambda-j.

5. As we show in the figure, we grow an interval around the **graph** eigenvalue lambda_n-j. The interval grows until it hits the first **graphon** eigenvalue lambda-**i** that is not lambda_**j**. We call this margin d_2

6. The minimum of the margins d_1 and d_2 is the eigenvalue margin for d of lambda_j comma lambda_n-j. This margin determines the convergence properties of the eigenvector associated with lambda_n-j to the eigenvector associated with lambda_j.

Slide 41: Convergence of Eigenvectors

1. We are now ready to state a classical theorem for the convergence of eigenfunctions of linear operators.

2. Given a graphon W and a graphon W_G-n induced by the graph G_n

3. We consider the graphon eigenvalue lambda_j along with the graph eigenvalue lambda_n-j. The latter is also an eigenvalue of the induced graphon.

4. Then, the distance between the corresponding associated eigenfunctions. Namely, the eigenfunction phi_j of the graphon W and the eigenfunction phi_n-j of the induced graphon W_G_n.

5. Is bounded by the product of pi over 2

6. With the ratio between the norm of the difference between the graphon W and the induced graph W_G_n

7. And the eigenvalue margin for the eigenvalues lambda_j and lambda_n-j.

8. The theorem implies convergence of the eigenfunctions for as long as the graph sequence converges to the graph. Indeed, if the graph sequence converges to the graphon, the induced graphon sequence converges to the graphon. The norm in the right hand side of the inequality vanishes and it therefore must be that the eigenfunctions converge.

9. This is true irrespectively of the eigenvalue margin. However, as the eigenvalue margin decreases, it takes a smaller value of the norm of the difference between the graphon and the induced graph to cancel it out. This means that we need to go deeper into the sequence index n to claim convergence. All eigenvectors converge. But it takes larger graphs for convergence to manifest when the eigenvalue margin is small.

Slide 42: The GFT Does Not Converge to the WFT

1. Herein lies the reason why the GFT does not converge to the WFT. For graphon eigenvalues close to 0 the eigenvalue margin vanishes.

2. These has to be the case because we know there are an infinite number of eigenvalues in the the minus-c-c interval. Eigenvalues accumulate at 0.

3. Thus, for any fixed sequence index n and constant epsilon, we have some eigenvalue index j for which the right hand side of the bound in the previous theorem exceeded epsilon. We just have to move the eigenvalue sufficiently close to zero so that the eigenvalue margin becomes sufficiently small. Remember that we have fixed the index n.

4. This the opposite of what we need for a convergence claim.

5. Which would be that for all constants epsilon, **all iteration indexes n** that exceed a certain n_zero **and all eigenvalue indexes j**, we can claim the bound to be smaller than epsilon. For those of you that know the term, we can claim convergence of individual eigenvectors. But we cannot claim **uniform** convergence of the set of eigenvectors.

Slide 43: Graphon Bandlimited Signals

1. The resolution of this problem is quite simple. However disappointing. We restrict attention to graphon bandlimited signals. Which are those whose WFT components are null below a certain threshold.

2. Formally, A graphon signal W-X is said to be c-bandlimited.

3. For some strictly positive bandwidth c.

4. If the Fourier coefficients of the graphon Fourier transform associated to eigenvalues whose absolute values is less than c is 0. We just require the WFT components associated with eigenvalues between -c and c to be null.

Slide 44: Bandlimited and Not-Bandlimited Graphon Signals

1. This definition is not simple. To emphasize this point suppose that we have a signal that is not graphon-bandlimited as we show in the figure.

2. We can make it graphon c-bandlimited by nullifying all of the WFT components of the signal that lie between minus-c and c.

Slide 45: Graph Fourier Transform Convergence for Bandlimited Signals

1. Introducing a bandwidth limit eliminates the problems associated with the accumulation of eigenvalues around zero. We eliminate the challenge of having garçon eigenvalues too close to each other. It is therefore not difficult to see that we can claim convergence of the sequence of **graph** Fourier transforms to the **graphon** Fourier transform.

2. Consider then a sequence of graph signals G_n-x_n. These signals converge to the graphon signal W-X, which we assume it is c-bandlimited. Observe ow the bandlimited assumption is on the graphon signal.

3. We then have that there exist a sequence of graph permutations pi_n such that

4. The sequence of GFTs of the graph signals G_n-x_n converges to the WFT of the graphon signal W-X.

5. The proof of this result is not difficult. We have eliminated the challenge that arises because of the accumulation of eigenvalues at 0. We are providing the proof in a supplementary material that you can download from the course webpage.

Slide 46: Inverse Graph Fourier Transform Convergence for Bandlimited Signals

1. The same can be claimed of the inverse GFT. Namely that the sequence of inverse **graph** Fourier transforms converges to the inverse **graphon** Fourier transform.

2. We consider then a sequence of graph Fourier transforms G_n-x_n-hat that converge to a graphon Fourier transform W-X. This requires convergence of the graph sequence to the graphon and convergence of the sequence of GFTs to the WFT.

3. The graphon Fourier transform is assumed to be associated with a signal that is c-bandlimited.

4. We then have that there exist a sequence of graph permutations pi_n such that

5. The sequence of inverse **graph** Fourier transforms converges to the inverse **graphon** Fourier transform.

6. The proof of this result is also not difficult. You can find it in supplementary materials that you can download from the course webpage.

Slide 47: Graph Fourier Transform Convergence for Bandlimited Signals

1. The convergence of the GFT sequence to the WFT depends on the convergence of graph eigenvalues to graphon eigenvalues.

2. We know that graph eigenvalues approach graphon eigenvalues as n grows. This bodes well for convergence of the GFT. This is mostly true. The convergence of eigenvalues implies convergence of the GFT. But this is only part of the story. We have a technical complication.

Slide 48: Graph Fourier Transform Convergence for Bandlimited Signals

1. As the eigenvalue index grows, the eigenvalues of the graph and the graphon become difficult to tell apart as they accumulated around 0. This precludes uniform convergence of eigenvectors.

2. And leads to a GFT convergence result that applies to graphon bandlimited signasl. Those that don't have GFT components associated with eigenvalues that lie below a threshold c in absolute value. At this point we must recognize that out interest in the GFT

and the WFT lies in their ability to explain the behavior of graph filters. We work on this next.

9.6 Graphon Filters

Slide 49: Graphon Filters - Title Page

1. We give the definition of graphon filters and give their frequency response. We further show that the frequency response is independent of the graphon itself.

Slide 50: Graphon Filters

1. Every graphon W induces a graphon shift operator T_W. By recursively applying the graphon shift operator to a graphon signal, we can create a diffusion sequence
2. With the k th element of that diffusion sequence, T_w^k X,
3. Given by the integral, evaluated between 0 and 1,
4. Of the product between the graphon W
5. And the previous element in the diffusion sequence, T_w^(k - 1) of X
6. For this definition to be complete, the initial term of the diffusion sequence has to be defined as the signal X itself.
7. We can now use that diffusion sequence induced by a graphon shift operator to define a graphon filter. Formally, we say that a graphon filter of order K is defined by the filter coefficients h_k and produces outputs Y given by the sum —- up to order K —- of the product between the coefficient h_k and the k-th element of the graphon diffusion sequence T_w^k of X
8. A graphon filter maps a graphon signal X to another graphon signal Y. Relying on our standard notation for operators, we then denote a graphon filter by T_H-X.
9. Note that a graphon filter is nothing more than a linear combination.

10. Of successive graphon diffusions of the original signal.

11. Modulated by the corresponding filter coefficients h_k

Slide 51: Graphon Filters and Graph Filters

1. In this sense, a graphon filter has the same algebraic structure of a graph filter. As we saw in previous lectures, a graph filter can be defined as a polynomial on the graph shift operator. That is, a graph filter of order K is comprised of successive graph diffusions (up to order K) modulated by the corresponding filter taps. Likewise a graphon filter of order K is comprised of successive graphon diffusions —- represented by recursive applications of the graphon shift operator T_w —- also modulated by filter coefficients h_k.

2. The only difference lies in the shift operator. Instead of a graph shift operator, we now rely on the integral graphon shift operator T_w of X instead. But, otherwise, we still encounter the familiar shift register structure. With its familiar use of shift, scale, and sum operations.

3. Indeed, to construct the output of a graphon filter, we start with the input signal X

4. Which, modulated by the initial filter coefficient h_0,

5. Is the first element that we add to construct the output

6. We then apply the graphon shift operator T_w to the input signal, observing the first diffusion of the graphon signal/

7. This quantity is scaled by the corresponding filter tap, h_1

8. And summed towards the output.

9. With another application of the graphon shift operator, we now obtain the third element in the diffusion sequence.

10. Which is modulated by the corresponding filter coefficient h_2

11. And summed towards the output.

12. By yet another application of the graphon shift operator, we now observe the fourth element in the diffusion sequence

13. Which is then modulated by the corresponding filter coefficient h_3

14. And added to the output.

15. Since we have a filter of order 4 this is the output of the graphon filter. The process is the same for higher order graphon filters. We just repeat the shifting, the scaling, and the summing a few more times.

Slide 52: Graphon filters in the Graphon Fourier Transform Domain

1. Now, we will leverage the graphon Fourier transform to analyze graphon filters in the frequency domain. First, we compute the graphon Fourier transform of the input signal, X hat, with each component X hat_j given by the integral of the product between the signal original signal X and the j-th eigenfunction, phi j,

2. And the graphon Fourier transform of the output, Y hat, defined in a similar manner

3. Now, we state a theorem for the graph frequency representation of graphon filters

4. According to which, given a graphon filter T_H with coefficients h_k,

5. The components of the graphon Fourier transforms of the input and output signals are related by

6. That is, the j_th component of the graphon Fourier transform of the output, Y hat_j, is equal to the product of the j_th component of the graphon Fourier transform of the input signal, X hat_j, and a polynomial on the j_th eigenvalue of the graphon shift operator, T_w, modulated by the filter coefficients h_k

7. That is the same polynomial that defines the graphon filter, but with the eigenvalue lambda j —- and not the graphon shift operator —- as a variable

Slide 53: Graphon Frequency Response

1. Since the j_th component of the graphon Fourier transform, Y hat _j, depends only on the j_th component of the graphon Fourier transform of the input signal, X hat_j, and a

polynomial on the j_th eigenvalue, we can conclude that graphon filters are pointwise in the graphon fourier transform domain. In the same way that graph filters are pointwise in the graph domain.

2. That pointwise characteristic of graphon filters in the graphon Fourier transform will allows us to define the frequency response of a graphon filter

3. Given a graphon filter with coefficients h,

4. the frequency response of the graphon filter is defined as a polynomial on a scalar variable lambda modulated by coefficients h_k.

5. This definition is such that we can write the output of a graphon filter in the graphon Fourier transform domain by multiplying the input GFT components X_j by h-of-lambda_j.

6. But a very important observation is that this is exactly the same definition of the frequency response of a graph filter with the same filter coefficients. This is so important that I will repeat it. This is the exact same definition of the frequency response of a graph filter. It doesn’t matter that we are now working with graphons.

Slide 54: Frequency Response is Independent of the Graphon

1. Why is this so important? Well, since the frequency response of a graphon filter and a graph filter are the same if they share the same filter coefficients, this function is a representation of both.

2. The effect of a graphon filter is to instantiate graphon eigenvalues.

3. And the effect of a graph filter is to instantiate graph eigenvalues.

4. If we now consider a graph sequence converging to a graphon, we know that the eigenvalues converge. This must mean that:

5. The filter **transfers**! From the graph to graphon. Fom the graphon to the graph. Across different graphs drawn from the graphon. Even if their numbers of nodes are different. This is the basis for our analysis of the transferability of graph filters and GNNs that we will undertake next week.