#	Question	Answer(s)
1	will these last 2 lectures be on the final?	All lectures, homeworks, and discussions are in scope for the final
2	why was the previous idea of plotting error vs degree not the best?	For the training data error usually is decreasing when degree is increasing, but that sometimes is simply due to we are fitting to noise.
3	are cost, error, and noise all the same thing?	Cost and error are similar ideas. Noise is something that causes error, but the actual error is dependent on the model we create (e.g. the degree of our polynomial in this case).
4	What's the difference between validation and test?	validation data help us determine some parameters in the model (for example the degree of polynomial). After we finalize our model, we will use test data to evaluate the model.
5	why are the cost vs degree graphs inverses of one another when we use all the data vs just the small subset?/	The small subset gives us less noise overall, so we are not as affected by noise in the model.
6	How does the error increased as degree increased?	As the degree increases, the model will try to fit more closely to whatever noise we have. When the data set is small, the noise "average" is not as good, so we may be overfitting to the noise.
7	So validation is like a pre-test test for our model?	usually at the validation step our model is not finalized -- we are still tuning the model. At the test step, we should no longer change the model -- if the test result is not ideal, we need to start over with a different model and data, instead of fine-tuning the model to improve the test result.
8	howcome the second cost vs degree curve didn't look the same, wouldn't higher order polynomials just fit the noise of the test data? i.e I don't see how it got worse?	In the 2nd case, we used a small subset of the complete data for developing the model. This usually means we are less affected by noise overall, i.e. even in higher order polynomials, there is less noise for the data to fit to. when we compare to the real data, we get a better "average" cost at lower order polynomials because it tracks the "real" model better rather than tracking the noise.
9	so in the real world, do all devices have the satellite songs already pre-installed to be able to use GPS? or do different companies have access only to their own satellites?	There are multiple satellite networks, run by different countries / agencies. Depending on which network your device uses, your devices has to already know the songs from those satellites.
10	What are the weights?	The weights are the -1 , 2, and 0.5 above. These can occur from signal amplitude loss over a long distance, or if the satellites are programmed to transmit using larger or small amplitudes
11	do our phones just ignore the songs of sattellites that are not part of their network, or do satellites just not transmit to specific devices?	Different networks transmit at different radio frequency bands, so our phones will ignore the songs of sattellites that are not part of their network
12	is shift the same as offset?	Yes
13	why isn't the noise scaled by alpha?	We're not looking for the alpha that the noise is scaled by, so we don't really care about the value. Its just there, and we need to account for it.
14	what's the n vector here again? I'm not quite getting it	n is the noise in the receiver.
15	have we talked about what constitutes a "good spike" or is it all relative?	We haven't talked about it yet. It will come soon in the OMP content.
16	why does a huge negative spike show a high cross correlation at that point	Think geometrically. The negative spike means anti-similarity, that its similar but in the opposite direction. For these satellite applications, its still useful
17	how does a negative spike prove that a satellite is on? don't high positive values of the cross-correlation indicate the satellite is on?	A sharp negative spike means we are anti-similar, so its similar but in the opposite direction. this means we have an alpha coefficient of -1
18	if there are no peaks, what does that mean? and what about in the presence of noise?	If there's no peaks, that means we did not observe the song / satellite in the samples we took. When there is noise present, it can mean that the noise is "louder" than the song (we haven't discussed this intuition)
19	how do we get the position of satellites, the process described so far doesn't mention it	Right now we're talking about finding the shift. Once we know the shift, we can compute distance, and from distance, we can compute position
20	is grass the same thing as noise?	Not necessarily. Grass can be due to our 'songs' are not prefectly orthogonal to each other or to itself at different shifts, so there will be some non-zero number in the cross-correlations
21	do we know the noise?	Somewhat. Noise is random, but we can learn a lot about the characteristics of the noise. Take CS70 and EE126 to learn more about how we deal with noise.
22	aren't we solving for both alphas and k's?	Yes we are solving for both alphas and k's
23	so our r is going to be 48 entries long? or do we need to find 48 unknowns with only 24 entries of r	r will be long, and we want to formulate the 24+ equations to solve for the 24 unknowns
24	when prof says that there are 24 satellites, is that in the whole world? or what satellites is she referring to?	You can think that there are totally 24 of them around the Earth, although actually there are a few more for backup.
25	Can you explain the difference between OMP and what we have already learned?	OMP is going to be a specific algorithm / technique for solving for shifts. Previously, we did this just by looking at the cross correlation of vec(r) with each song, but we're going to do some more advanced techniques to speed up the search

26	I thought we were told in discusssion that we have to shift from -(len of r) to +(len of s) not just the length of s	what we said in discussion is the more general form and is correct. Looking at just within the length of s is okay if we are restricting ourselves to that region, which we may do depending on the song design.
27	On the test, should we find the correlation of negative shifts as well?	It depends on the problem, so read the problem and see if it's possible to have a negative peaks. But if there's indeed a significant negative peak it probably means some signal.
28	Do we calculate cross correlation starting from a shift of 0. In last week's discussion, we used negative shifts in addition to using a shift of 0	The versoin from last week's discussion is more generally correct. Depending on the context or song design, we may consider only nonnegative shifts to be "valid".
29	how do we know cross correlation works for the superposition of signals (the received signal being the one with superposition)? haven't we only looked at two signals so far, both of which were not a combination of multiple signals?	We can always take cross correlation between any two vectors. In this context, we want all the songs to be distinct or orthogonal not just to each other, but to all shifts of. Then if we take cross correlation between each song, it will not give us any peaks
30	why are we cross correlating with y and not r hat, r hat is an approximation of r right?	rhat would only contain information from s1 and s2, so we would have no correlation with s3. We have to use y if we want the correlation with s3
31	Why do we still check the cross correlation of the residual with S2 if we already predicted the weight of it in the previous iteration?	we adjust our alpha value based on the "next" satellite found. Because S1 and S2 are not completely orthogonal, doing a projection into that space (which is getting rhat) requires doing the LS solution using both s1 and s2
32	whats residual (y)? and what are we doing with the are shift and lag the same thing here?	shift and lag are the same thing. residual(y) is "remaining" in the signal after we have found the "amount" of s1 and s2
33	What does a weight represent again?	It represents how 'loud' is the song in the received signal.
34	is what we're doing right now OMP? like is this approximation process called OMP	yes, this iterative process of checking for more and more songs is OMP
35	so residual is a way to measure error?	residual is the remaining signal after we remove the known songs from the received signal
36	When do we stop iteration	Either when we run out of songs or our residual falls below some pre-determined residual value.
37	How many iteration should we do?	We iterate until either when we run out of songs or our residual falls below some pre-determined residual value.
38	why are we using residuals in the cross correlation?	We're looking in what remains in the original signal that we haven't extracted out already. We already extracted out all the information related to s1 and s2
39	Does the first calculation of weight usually give us a good approximation. I dont understand how we can get a good approximate if we are assuming our R vector only has 1 satellite and we calculate our first weight based on that	We know the 'songs' are almost orthogonal to each other, so we don't need to assume there is only 1 satellite in the received signal
40	if there was hypothetically no noise, should the peaks of all signals be the same value?	Not necessarily. There could still be signal loss from transmitting over a long distance or if the satellites were transmitting with different powers. But if there were no loss and they all started with the same power, then yes
41	would we have to continue the process if alpha3 was off by a lot?	We can stop when either we run out of songs or our residual falls below some pre-determined residual value.
42	So is omp not in scope?	
43	When and why should we use OMP?	live answered
44	Can we optimize this algorithm?	"optimize" is a broad term and depends on the context. But in this case, we can speed up the algorithm if we design (and assume) the songs to be orthogonal or almost orthogonal
45	was that the last slide?	live answered
46	So we are not being tested on OMP?	
47	Will Professor Waller teach any course next year ? Doesn't seem like it according to the current tentative schedule, but wondering if a course will be added	live answered
48	why are we correlating the songs with themselves again? and how does this tell us if they're on?	This is telling us about the similarity of each signal. We want to the songs to be orthogonal, i.e. completely distinct from each other
49	Where can I find a mechanical example of OMP?	live answered
50	will professor waller be teaching 118 in the near future?	live answered
51	Can I take 127 and 189 ast the same semester? Will it be hard?	Depending on who you ask and who is teaching, 127 is a useful prereq for 189. I would say that combination will take up a lot of your time.
52	what part of this class is the most useful for data analytics	Least squares and regression :)