#	Question	Answer(s)
1	In the proof, we said e is perp to col(A). Why didn't we only consider e being perp to bhat?	We want to find b_hat that gives us the smallest error. Its not so much that we don't consider just bhat, but we need to find bhat in col(A).
2	Why is the equation Ax = b + e, and not Ax = b - e, since b^ = b - e in the diagram?	e can be a neg number
3	Why must A be transposed in order for it to be orthagonal to the error vector?	live answered
4	we're always going to know the A and b right	Yes, and we're trying to find x. Sort of like our imaging problems
5	how did we minimize e in the derivation again?	live answered
6	What is the geometric meaning of a transpose	There is an interpretation with orthogonal subspaces
7	Is this an accurate graph? If I cry more will I get a better score?	#memes
8	Will I get a higher score on my exam if I cry more. If so I can totally do that	#memes...for real though, we don't have real data on this. Study hard, come to office hours, ask questions.
9	Why is the slope c instead of b?	The intercept is c. We're using another variable since we've been using the "b" vector a lot already.
10	is [m c] supposed to be [m b]	The intercept is c. We're using another variable since we've been using the "b" vector a lot already.
11	Why is the unknown vector [m c] and not [m b]?	The intercept is c. We're using another variable since we've been using the "b" vector a lot already.
12	Why is the error the vertical distance and not the perpindicular distance to the line/subspace, like before?	We will be minimizing the perpendicular / square distance. We're just showing some distance to the line to visualize
13	So is each row in A like an individual point on the graph?	yup!
14	Visually, when we took the Col(A) before, we could interpret it as the "ideal" subspace? In this case, since A contains our raw x's which aren't perfectly aligned, what does Col(A) represent visually?	Our two vectors are 1s and a linear term; so the space will be all linear "lines" (with offset)
15	will we get iPython on the final	highly unlikely
16	what do the degrees mean	Its the degrees of the polynomial we're fitting. e.g 1 -> linear; 2 -> quadratic; 3 -> cubic, etc.
17	when we go to higher degrees we are no longer linear right	correct. degree =1 is linear
18	Won't it overfit at some point	live answered
19	How would you know the underlying model?	live answered
20	What does the meme mean?	live answered
21	Shouldn't the first entry be x^2 is opposed to x1^2?	The left most column is the x^2. we make that vector from squaring each of the x_i values.
22	if one of the coefficients were squared we wouldn't be able to do this right	live answered
23	if we're kind of behind on these topics, how would you recommend catching up?	Start with reviewing the notes and discussions. Then come to office hours to fill in gaps.
24	should there be a constant term for the polynomial?	For the ellipse, no. We can absorb the constant term into the right hand side
25	not regarding the lecture but does the final have similar number of questions as the midterm (so around the same number of points but more time)?	You'll have roughly the same time per question as with the midterms
26	trivial means that the only solution is for the x vector to be the zero vector right	Yes, the only solution to Ax=0 is x=0.
27	aren't we missing the square root?	technically yes, norm = sqrt(inner product). But since norm is always positive, then inner product is always positive. so proving it for the inner product is sufficient
28	Don't we also need to show that every vector in Null (A) is also in the transposed null space	We are showing Null(A^T A) = Null(A), so we don't need to show anything about Null(A^T).
29	how do we take the norm of a vector, or the inner product of a matrix	norm of a vector is the square root of its inner product (with itself). We only define the inner product of two vectors (nx1 matrix).
30	Isn't the first term v transpose?	live answered