Lecture #15: Background

So far in this course, we have concentrated on the forward problem, i.e., given an acoustic environment (the ocean and seabed) and some physics/wave theory (the wave equation in its various forms and approximations), we calculate an acoustic field (pressure, particle displacement/velocity) or some derivative quantity (travel time is a common one). In general form, this can be written as G(m)=d, where G is the physics equation, m is the environment (model), and d is the acoustic field quantity measured (data). This form is much more general than just acoustics, but that's not our worry (yet). (The linear form is Gm=d, and is "matrix friendly.")

In the grand scheme of things, we know the acoustic wave equation very well, and we can measure pressure/velocity/travel time/acoustic stuff very precisely – but what about the environment? That's the kicker. The ocean is turbulent, anisotropic, inhomogeneous, non-stationary, full of noisy biota (us included), has a complicated seabed under it and – well, its messy! In doing the forward problem, we've ignored this and blithely put in a "known" ocean. But that's not real, and in fact, our environmental inputs to the forward problem are sometimes quite poor. So what do you do? Well, the obvious answer is "you measure the ocean better!" And that's where inverse theory comes in. If we invert our little equation above to $m = G^{-1}d$, we can now use our acoustic measurement as data (d was chosen with malice aforethought) to estimate an ocean/seabed model. As mentioned, the d=Gm form is very general, and will allow us to throw in all sorts of data types, as well as a priori information and dynamical models, into the mix.

Turning to the notes, ABT start Chapter 1 with a nice introductory discussion of the various flavors of inverse theory – inverse, parameter estimation, discrete versus continuous, linear and nonlinear, system identification, etc. One quickly gets the idea this is a big field (that, alas, we again will treat all too briefly). The forms the equations can take are briefly covered, and two forms in particular, the Fredholm integral equation of the first kind (IFK) and the Fourier integral form, are pointed out as especially important.

The notes then look at some "inverse forms" that we already have seen in the course of the forward problem and instrumentation discussions. These include: 1) the form relating Δc to the modal wavenumber perturbations, Δk_n, 2) the Fourier integral form for bottom roughness, and 3) the travel time perturbation form seen for tomography. All these have IFK or Fourier integral forms, so they are well posed candidates for the inverse problem! The ocean tomography problem gets a little extra play here, in that this problem leads to a 3D description of the ocean that will be useful for acoustics and other applications. We get at that 3D description as follows. The ocean has a rather specific variety of vertical structures it will allow, which are discussed in section 18.6 of the Spindel chapter of Medwin (Ed), and in particular in Eqs. 18.16 and 18.17. These structures are the vertical modes of the ocean variability about the mean, and can be

described best by either ocean dynamic modes (very like the IW modes we saw in the notes, only lower frequency) or most commonly by "empirical orthogonal function (EOF)" modes which we will treat next. Before doing that, we note that by using vertical modes, we "integrate out" the z-direction part of the problem into a simpler set of x-y inverse problems for each mode. That is quite useful, especially since the ocean has a red spectrum that you usually only need a few vertical modes to describe!

The EOF's are simply the eigenvalues and eigenvectors of the covariance matrix, and the next section of the notes shows a very simple example of how to obtain these for a typical set of ocean measurements. (Specifically, it discusses a set of CTD casts distilled down to soundspeed – the well beloved "poking holes in the ocean" measurement that all seagoing oceanographers make). The notes are clear here, I think, so what's the next step?

If I have the EOF modes and their coefficients at the measurement sites, I have a way to very nicely estimate the 3-D ocean everywhere else! Specifically, we interpolate the $a_i(x, y)$ EOF modal coefficients in 2-D for each mode. This gives a smooth, consistent interpolation, and methods to do this are described in the notes. There are caveats in doing this also (e.g., respecting the space/time correlation lengths of the measurements, watching azimuthal symmetry assumptions, etc.), but by and large these are just common sense. A few examples of using EOF's are shown in the notes - interpolating the Gulf Stream and obtaining a 3D ocean model in the vicinity of an ocean seamount!

In section 1.4 of ABT, there is a brief, but very useful, discussion of "why inverse problems are hard. This is a great"in nucis" prelude to the rest of the book (and this segment of the class).

The next section of the notes is from an older version of Aki and Richards "Quantitative seismology" specifically Chapter 12 of volume 2. Aki and Richards (AR) were ABT's "mentors" (see their preface), and indeed the ABT book can be taken as the modern expansion of AR Chapter 12.

In Chapter 12, AR show the forward problem, the inverse, the resolution matrix, and the variance matrix right at the start, equations 12.79 through 12.86. This "answer" (the inverse) and its "error bars" (resolution and variance) are what we want to get at in a formal way, and that is the meat of the rest that follows. Before getting into the more detailed linear algebra, however, AR pose a tiny system of three inconsistent simultaneous equations, which provide a remarkably nice "toy problem" (eqs. 12.87) to illustrate all the theory concepts developed! (I am a large fan of such toy models, as you've seen!).

The next few pages of AR after Eqs. 12.87 are devoted to developing the Lanczos decomposition (also called the singular value decomposition or SVD), and emphasizing the data and model spaces. This is also done in the first chapters of ABT, and the developments are similar. With the SVD in hand, one can obtain the generalized inverse, and also the resolution and variance, very conveniently. AR do this very compactly, whereas ABT has a bit longer, more expansive discussion. It is well worth reading both.

The "star of the show" in AR, however, is the toy model. Using it, AR give very clear, clean, analytic examples of all the linear algebra concepts developed. This is why I keep coming back to AR as a resource.

Before jumping to the next part, let me just stress something that AR probably thought was too obvious – that linear inverse theory can be used (and should be used, when possible) as a great tool for experimental design!! How the data illuminates the model space (or fails to), the size of the features it resolves, and the model error one can expect all tell you about the data set you should collect, and need to collect. Moreover, the inverse scheme accommodates all sorts of data types (as we'll see), so one can choose among measurement options to get an optimal observing system in terms of data adequacy, cost, ease of performance, and anything else you want to drop into an optimization cost function! This is the "formal mathematical" version of what I talked about in my shallow water acoustics lecture!

Now we turn to ABT, and its development of the SVD and generalized inverse. Rather than derive the SVD methodically and then talk about its uses, as AR do, ABT just state it up front, and then describe its uses in a bit more detail than AR. Thus, AR and ABT provide complementary material on the SVD.

After writing down the SVD, ABT show the generalized inverse (Moore-Penrose pseudo inverse) solution in Section 3.1. The discussion after this, Section 3.2, on the "co-variance and resolution of the generalized inverse" is similar to AR, but with a bit more discussion. The next part, Section 3.3, of ABT, on the stability of the inverse and condition number, is not covered very much in AR, and this is an important section. The discrete Picard condition for stability, and fixing instability via the truncated SVD (TSVD) are discussed, which directly leads into the general topic of regularization. A "rank deficient" ($p=\text{rank}(G)<n$) toy tomography problem is shown in Section 3.4, using the SVD. Note that ABT use small, MATLAB solvable, toy systems for examples, as opposed to AR's very simple analytical system. This is the "modern era" toy problem, where one can kick out some numerical results quickly and easily. Section 3.5 on "discrete ill posed problems" poses another potential pitfall in inverse problems – over discretization of the Fredholm integral. (One can get greedy and add more layers/pixels, but you still can't beat the overall resolution limit, and you can actually cause trouble for yourself, as noted above!)

Chapter four of ABT covers the important topic of Tikhonov regularization, which gives a way to "optimally" treat the issues of instability and ill-posedness. This is done by minimizing the solution norm or the data misfit, or a combination of them. This gives a "trade-off curve" between the solution norm and data misfit, which often has an L-shape to it. Picking the "knee" of this curve is one common optimal solution criterion.

Section 4.2 goes through the important topic of damped least squares in some detail, which is perhaps the "canonical example" of regularization. Section 4.4 on ABT covers "higher order regularization," which allows us to impose some prejudices on the solution (roughness, smoothness) by minimizing Lm instead of m, where L numerically inserts our nasssty little prejudicesss. Gollum.

The first four chapters of ABT cover the basics of linear inverse theory, including regularization, but there are a few other bits and pieces that should be mentioned before leaving. A brief section from Mencke on the "resolution and variance trade-off" is worth reading, as an additional note on that topic. The ABT Chapter 7.1 material on non-negative least squares is also important, as many real world problems can only have positive definite solutions. An example of particle size distribution estimation using multi frequency acoustics is presented – the particle radii must be positive!

Appendix A from Mencke on "Implementing Constraints with Laguange Multipliers" is a rather short, (to my view, too short) introduction to the very important topic of how one incorporates "additional information and constraints" into a basic inverse problem. Both in Mencke and ABT, there are decent derivations, but not really good examples. This is maybe a good homework problem for me (and you!).

Finally in Lecture 15, I show section 7.3 of the KPL book, which shows one application of the generalized inverse to some nice bottom inverse data collected off Corpus Christi, Texas. The bottom line results are the compressional wave speed profiles in the bottom and the associated "resolution length" vs. depth in the bottom. These results were "simple" to obtain from the modal eigenvalue perturbations we discussed, but also agreed well with broadband and fully nonlinear acoustic inverses! So this "simple" theory kinda works well at times!

2.682 Acoustical Oceanography

Spring 2012