

# The Oceanic Inverse Problem

**Overdetermined and underdetermined problems** The ATSR example had exactly the same number of observations as quantities to be estimated. This unusual, and also undesirable: why?

As a general rule, we would like our inverse problems to be *overdetermined*, meaning the matrix  $\mathbf{K}$  in

$$\mathbf{y} = \mathbf{K}\mathbf{x} + \boldsymbol{\epsilon} \quad (1)$$

has more rows than columns. We've already shown that the variance in the Gauss-Markov optimal estimate is given by:

$$V(\tilde{\mathbf{x}}) = (\mathbf{K}^T \mathbf{S}_\epsilon \mathbf{K} + \mathbf{S}_a^{-1})^{-1}, \quad (2)$$

so for any given noise variance (size of diagonal elements of  $\mathbf{S}_\epsilon$ ), the first term increases the larger the rank of  $\mathbf{S}_\epsilon$ , so the estimate depends more on the data and less on the *a priori* (see problems).

Scales of motion in the ocean are typically much smaller than the atmosphere (why?) and observations are much sparser. This means that any attempt to characterise the flow in the oceans' interior is likely to be *undetermined*, meaning the matrix  $\mathbf{K}$  in

$$\mathbf{y} = \mathbf{K}\mathbf{x} + \boldsymbol{\epsilon} \quad (3)$$

has more columns than rows.

---

**Inferring ocean currents from hydrography** The simplest observations to make in the ocean are profiles of temperature and salinity as a function of pressure. It's a lot harder to measure  $T$  and  $S$  as a function of absolute depth: why? What does "absolute depth" mean in the ocean anyway?

Over most of the oceans and most of the water column, motions are almost exactly hydrostatic:

$$\frac{\partial p}{\partial z} = -\rho g \quad (4)$$

and fairly close to geostrophic

$$\frac{\partial p}{\partial y} = -\rho_0 f u \quad (5)$$

$$\frac{\partial p}{\partial x} = \rho_0 f v, \quad (6)$$

where  $\rho_0$  is the characteristic density of sea water. Combining these gives the thermal wind equations

$$\frac{g}{\rho_0} \frac{\partial \rho}{\partial y} = f \frac{\partial u}{\partial z} \quad (7)$$

$$\frac{g}{\rho_0} \frac{\partial \rho}{\partial x} = -f \frac{\partial v}{\partial z}. \quad (8)$$

So given a set of hydrographic profiles, we can work out geostrophic currents as function of depth to within an integration constant:

$$u_G(z) = \frac{g}{\rho_0 f} \int_{z'=z_R}^z \frac{\partial \rho}{\partial y} dz' + u_R \quad (9)$$

$$\rho_0 v_G(z) = -\frac{g}{\rho_0 f} \int_{z'=z_R}^z \frac{\partial \rho}{\partial x} dz' + v_R \quad (10)$$

But what are  $u_R$  and  $v_R$ , the “reference level flow” at  $z = z_R$ ?

---

**The reference level problem** If all we care about is flows near the surface, which are typically 10-100 times faster than abyssal flows, then assuming no motion below some depth such as 2000m is a reasonable approximation. But for problems like working out whether the North Atlantic Overturning Circulation is slowing down, this is insufficient. There have been numerous attempts to solve the reference level problem, starting with...

**Hidaka’s problem:** Hidaka (1949) proposed an approach to the reference level flow problem based on conservation of mass and salinity in a box with salinity observations at each corner. Assuming the scales of motion are larger than the dimensions of the box, he equated salinity flux between any two vertices with the average salinity along that line multiplied by the mass flux across that line. This apparently gives six equations with six unknowns, the mass flux across each wall and each diagonal. In fact, this system is *ill conditioned*, meaning there are fewer constraints than unknowns, as should be obvious from the fact that there are only three (relevant) independent observations (why not four)?

---

**The contribution of altimetry** Radar altimetry now allows variations in sea level to be measured from space with a precision of only a few centimetres. If  $\eta$  is the departure of the sea level from the geoid, then  $g/f \frac{\partial \eta}{\partial r}$ , where  $r$  is along-track distance, gives the across-track component of the near-surface geostrophic flow. If  $\frac{\partial \eta}{\partial r} > 0$ , which direction is the across-track flow in the Northern Hemisphere?

The problem now becomes measuring the geoid, which departs from a perfect oblate spheroid by hundreds of meters in places. Fortunately, the geoid doesn’t change very fast (timescales of mantle convection), so new precise gravimetric missions should allow us ultimately to solve the reference-level problem once-and-for-all by combining altimetry and hydrography.

See [oceanworld.tamu.edu/resources/ocng\\_textbook/chapter10/chapter10\\_03.htm](http://oceanworld.tamu.edu/resources/ocng_textbook/chapter10/chapter10_03.htm) for an excellent introduction.

---

**Baroclinic flows from altimetry** The simplest baroclinic model for the world oceans is the so-called “one-and-a-half-layer” model, in which a single, vertically homogenous active layer of warm (less-dense) sea-water overlying an inert, infinitely deep abyss. You’ll be encountering this model again in further GFD when we talk about the equatorial waveguide, but it provides a simple model to understand the use of altimetry. Some questions to think about before you attempt the problem (from Finals, 2005):

- How are the following quantities related to the slope of the sea surface relative to the geoid (work them out in this order): (a) pressure gradients in the abyss; (b) the slope of the thermocline; (c) pressure gradients in the active layer; (d) geostrophic currents in the active layer?
- Do geostrophic currents in the active layer depend on either the depth of the active layer or the density of the abyss? So, in this simple model, what depth-integrated quantity is proportional to the sea surface slope?
- And to tie this in with the previous lecture, if you make ten unbiased and independent observations of the same quantity, each with standard error  $\sigma$ , how would you estimate the true value of that quantity and what is the standard error of your estimate? Why? Can you suggest an alternative unbiased estimator of this quantity using only these ten observations? What is its standard error? Of what theorem is this a special case?
- If you make ten unbiased and independent observations of two quantities and find they have correlation  $\rho$ , what is the covariance matrix  $\mathbf{S}_\epsilon$  of your estimate of the resulting two-vector?

## Time-dependent inverse problems

So far, we have focussed on instantaneous, local inverse problems: estimating SST, geostrophic flow etc. This keeps things simple, but means we make no use of prior knowledge of the dynamics of the system we are studying. The rest of this lecture will focus on some of the creative things we can do when we bring together an understanding of optimal estimation theory with an understanding of the non-linear dynamics of the atmosphere-ocean system. This material is not examinable, but I'm including it to convince you that the theory we cover in remote sounding isn't just relevant to what sounds like a very practical engineering problem of working out the state of the atmosphere from a set of satellite radiances. It goes much deeper than that, because this is the basis of our understanding of how *information* is introduced, propagates and decays in a non-linear system.

We will cover:

- Optimal estimation applied to 2-D and 3-D problems: the optimal interpolation problem.
- Inferring initial conditions from observations taken at a later time: the adjoint problem.
- Singular vectors, error growth and adaptive observation systems.

We have learned that the Bayesian maximum-likelihood estimate of  $\mathbf{x}$  is

$$\tilde{\mathbf{x}} = (\mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{K} + \mathbf{S}_a^{-1})^{-1} (\mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{y} + \mathbf{S}_a^{-1} \mathbf{x}_a) \quad (11)$$

which can be rearranged (check!) to give

$$\tilde{\mathbf{x}} = \mathbf{x}_a + (\mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{K} + \mathbf{S}_a^{-1})^{-1} \mathbf{K}^T \mathbf{S}_\epsilon^{-1} (\mathbf{y} - \mathbf{K} \mathbf{x}_a) \quad (12)$$

$$\equiv \mathbf{x}_a + \mathbf{G} (\mathbf{y} - \mathbf{K} \mathbf{x}_a) \quad (13)$$

where  $\mathbf{G}$  is the Kalman Gain Matrix which operates on the data-*a priori* misfit term to produce the optimal increment on the “first guess”  $\mathbf{x}_a$ .

---

**If you’re keen on linear algebra** you can derive the alternative form of the Kalman gain matrix

$$\tilde{\mathbf{x}} = \mathbf{x}_a + \mathbf{S}_a \mathbf{K}^T (\mathbf{K} \mathbf{S}_a \mathbf{K}^T + \mathbf{S}_\epsilon)^{-1} (\mathbf{y} - \mathbf{K} \mathbf{x}_a) \quad (14)$$

Suppose  $\mathbf{S}_a$  has eigenvectors with zero eigenvalues: what would that mean in physical terms? Notice how this second form of  $\mathbf{G}$  makes clear how the increment can never have any component in these directions: why does this make sense from a physical point of view?

**Optimal interpolation** is the application of optimal estimation theory to generating continuous fields from incomplete and noisy observations. Suppose we have a set of noisy, gap-riddled observations of a process that we know, a priori, to be a sine wave. What is  $\mathbf{S}_a$  in this case? What does this tell us about the optimally interpolated solution?

---

**Introducing time evolution** Suppose we have a dynamical forward model (a weather model, for example) and observations come in continuously over some interval (a 12-hour assimilation cycle, for example), how do we find the initial conditions for the model that minimises the model-data misfit to all these observations? This is the classic problem in initialising weather forecasts. Let’s focus on the situation in which  $\mathbf{S}_a = \mathbf{I}_n$ ,  $\mathbf{S}_\epsilon = \mathbf{I}_m$  and  $\mathbf{x}_a = \mathbf{0}$ . The optimal estimator becomes:

$$\tilde{\mathbf{x}} = (\mathbf{K}^T \mathbf{K} + \mathbf{I})^{-1} \mathbf{K}^T \mathbf{y} \quad (15)$$

with variance  $(\mathbf{K}^T \mathbf{K} + \mathbf{I})^{-1}$ , where  $\mathbf{K}$  is the linear operator that predicts the observable quantity  $\mathbf{y}$  from the initial conditions  $\mathbf{x}$ . If  $\mathbf{x}$  and  $\mathbf{y}$  are a  $10^7$ -dimensional vectors, we can’t write down  $\mathbf{K}$  as a simple matrix, but we can express it as an operator, in terms of the weather model itself. But what about  $\mathbf{K}^T$ ?

---

**Adjoints:** Recall the vector differentiation result

$$\frac{\partial \mathbf{K} \mathbf{x}}{\partial \mathbf{x}} = \mathbf{K}^T \quad (16)$$

so  $\mathbf{K}^T$  is simply the linear sensitivity of the observations  $\mathbf{y}$  to the “control variables”  $\mathbf{x}$ . When we have a non-linear model, we need to linearise the equations of motion about a reference trajectory:  $\mathbf{K}^T$  then corresponds to the an operator which involves running a the (linearised) model backwards in time to compute the sensitivities of  $\mathbf{y}$  to  $\mathbf{x}$ ,  $\partial y_i / \partial x_j$ : this is known as “adjoint modelling”.

---

**Singular vectors:** If the observations provide a complete description of the state of the system at time  $t + \delta t$ , then  $\mathbf{K}$  is the linear propagator that takes a perturbation to the initial conditions,  $\mathbf{x}$ , into a change of state at the time of observation. Expressing  $\mathbf{K}$  in terms of its singular value decomposition  $\mathbf{K} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$ , where  $\mathbf{U}$  and  $\mathbf{V}$  are orthonormal matrices and  $\mathbf{\Lambda}$  is diagonal, provides vast amounts of information about the impact of observations and the evolution of errors.

The left singular vectors  $\mathbf{U}$  corresponding to the largest elements of  $\mathbf{\Lambda}$  are the dominant errors at time  $t + \delta t$ , while the right singular vectors  $\mathbf{V}$  correspond to the critical errors in initial conditions that evolve into these dominant errors.

Notice that

$$\mathbf{K}\mathbf{K}^T = \mathbf{U}\mathbf{\Lambda}^2\mathbf{U}^T \quad (17)$$

$$\mathbf{K}^T\mathbf{K} = \mathbf{V}\mathbf{\Lambda}^2\mathbf{V}^T \quad (18)$$

so the  $\mathbf{U}$  and  $\mathbf{V}$  are the eigenvectors of  $\mathbf{K}\mathbf{K}^T$  and  $\mathbf{K}^T\mathbf{K}$  respectively. This allows them to be calculated from the forward and adjoint linearised models even if we cannot hope to perform an explicit SVD of  $\mathbf{K}$ .

---

**But if we have a complete description of the state, why worry about what happened in the past?** If our description is noisy, we can exploit knowledge of the system dynamics encapsulated in  $\mathbf{K}$  to improve our optimal estimate: the variance in our optimal estimate of  $\mathbf{y}$ , viz.  $\tilde{\mathbf{y}} = \mathbf{K}\tilde{\mathbf{x}}$  is (recalling  $\mathbf{S}_a = \mathbf{I}_n$ ,  $\mathbf{S}_\epsilon = \mathbf{I}_m$ ):

$$\langle \tilde{\mathbf{y}}\tilde{\mathbf{y}}^T \rangle = \mathbf{K}(\mathbf{K}^T\mathbf{K} + \mathbf{I})^{-1}\mathbf{K}^T \quad (19)$$

$$= \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T\mathbf{V}(\mathbf{\Lambda}^2 + \mathbf{I})^{-1}\mathbf{V}^T\mathbf{V}\mathbf{\Lambda}\mathbf{U}^T \quad (20)$$

$$= \mathbf{U}\mathbf{\Lambda}^2(\mathbf{\Lambda}^2 + \mathbf{I})^{-1}\mathbf{U}^T \quad (21)$$

So the errors in each pattern  $\mathbf{u}_i$  vary as  $\lambda_i^2/(1 + \lambda_i^2)$ : if  $\lambda_i \ll 1$ , so perturbations contract, then the error in the optimal estimate is small in that direction.

---

**Adaptive observation systems:** If we have a limited number of observations to make at time  $t + \delta t$  then we minimise the uncertainty in  $\mathbf{x}$ ,

$$\langle \tilde{\mathbf{x}}\tilde{\mathbf{x}}^T \rangle = (\mathbf{K}^T\mathbf{K} + \mathbf{I})^{-1} = \mathbf{V}(\mathbf{\Lambda}^2 + \mathbf{I})^{-1}\mathbf{V}^T, \quad (22)$$

by aligning these observations with the dominant singular vectors of  $\mathbf{K}$ . Where these observations are required depends on the trajectory the equations have been linearised about: “flow-dependent” errors. This methodology is being used to target observations onto evolving storms as they develop in mid-ocean a few days before they arrive into inhabited regions.