Multiscale retrograde estimation and forecasting of chaotic nonlinear systems

Thomas Bewley, Flow Control Lab, Dept of MAE, UC San Diego, La Jolla CA 92093-0411

This talk will present a promising new strategy for the estimation and forecasting of chaotic nonlinear systems characterized by the related challenges of multiscale complexity and model uncertainty.

The phrase *model-based estimation* is often associated almost synonomously with the ubiquitous *Kalman filter*, which is in a certain sense optimal in the low-dimensional linear/quadratic/Gaussian (LQG) setting. The key principle upon which the Kalman filter is based is that, in a linear system, *Gaussian initial uncertainty of the state estimate and Gaussian disturbances and measurement noise result in Gaussian uncertainty of later optimal state estimates*. This important principle allows the Kalman filter to summarize all past system measurements with a single state estimate (of dimension N) and covariance estimate (of dimension N^2), the former representing the best estimate of the present system state based on all past measurements, and the later parameterizing completely the uncertainty of this estimate.

When the system of interest is an accurate (that is, high-dimensional, often with $N \gtrsim 10^6$) discretization of a multiscale infinite-dimensional system governed by a partial differential equation (PDE), the matrices at the heart of the Kalman filter approach are computationally intractable. The common remedy applied in this case is to follow an iterative vector-based approach based on repeated adjoint analysis of candidate system trajectories; when framed appropriately, this approach solves iteratively the same problem that the Kalman filter solves in a single shot, but with increased computational requirements (for complete convergence, which is often not required) and reduced storage requirements [$O(N)$ instead of $O(N^2)$]. This iterative approach is referred to by the weather forecasting community as *4Dvar*, and by the controls community as *moving-horizon estimation* (*MHE*). A hybrid approach which approximates the covariance matrices at the heart of the Kalman approach via the sum of the outer product of vectors is also available; this approach is referred to as a *reduced-rank Kalman filter*.

When the system of interest is not linear, a common “fix” is to linearize the system model (either about a specified trajectory of the system, or a representative mean state), to design a Kalman filter, then to throw the system nonlinearity back onto the estimator model at the eleventh hour, an algorithm referred to as the *extended Kalman filter*. Though the key principle mentioned above, upon which the Kalman filter (and related approaches) is based, is in fact true for infinitesimal uncertainties in smooth nonlinear systems, *this principle fails spectacularly in chaotic nonlinear systems for finite, yet still relatively small, disturbances typical in such systems* (see Figure 1). As anyone who has experienced the consequences of a bad weather forecast will readily attest, the perturbation of the state estimate from the actual system state is often not infinitesimal in the problem of weather forecasting. Indeed, problems of this class are so difficult just in the modeling and measuring of the system that perturbations of the state estimate from the actual system state are essentially guaranteed to be larger than anyone involved in the forecasting process would care to admit. The question addressed well by our new algorithm is how to deal with such finite perturbations, and the associated non-Gaussian distribution of uncertainty of the estimate, in a tractable manner.

Note that, in chaotic systems, the system trajectory moves on *attractor* in phase space, with the effects of the nonlinear terms non-negligible and in an averaged sense in some sort of balance with the effects of the linear terms in order to maintain the unsteady motion on the attractor. As the chaotic system moves on this attractor, the *Lyapunov exponent* measuring the time-averaged exponential rate of divergence of infinitesimally perturbed trajectories is positive, and the *local Lyapunov exponent* measuring the local exponential rate of divergence of infinitesimally perturbed trajectories is sometimes both positive and large. It is these “trouble spots” on the attractor that cause the most difficulties to existing state estimation algorithms.

The method that we propose in this talk (summarized briefly in Figure 2, Algorithm D) is unique from the standpoint that it *revisits past measurements in light of new data in an adjustable manner based on a quantitative measure of the quality of the current state estimate*. For more information, a GUI is available on the web implementing this method at `http://renaissance.ucsd.edu/retro`.

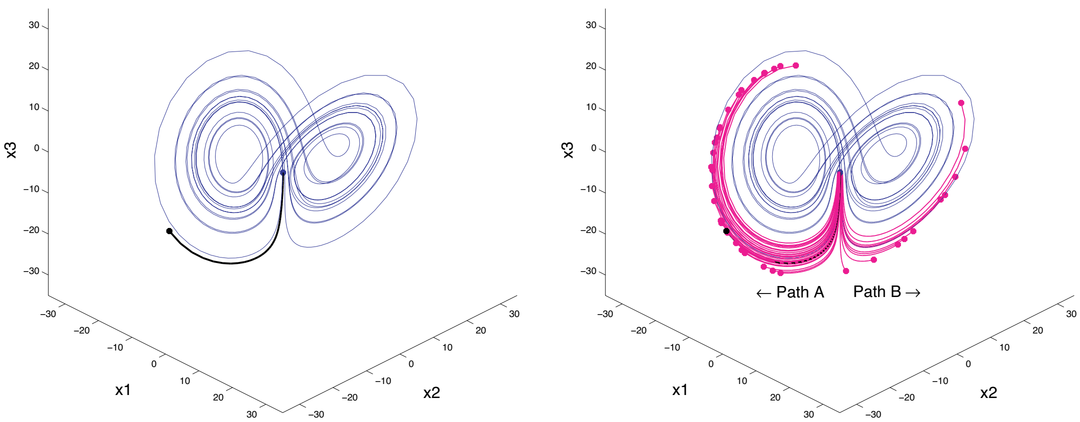

Figure 1: Demonstration of the nominal trajectory of the state estimate (left, black), and several perturbed trajectories of the state estimate (right, red) initiated at a point in the Lorenz system characterized by a large local Lyapunov exponent. In this test, the initial perturbations of the perturbed trajectories are very small (to plotting accuracy, they are on top of each other), and distributed in a Gaussian fashion. The final distribution of the perturbed trajectories, however, is highly non-Gaussian.

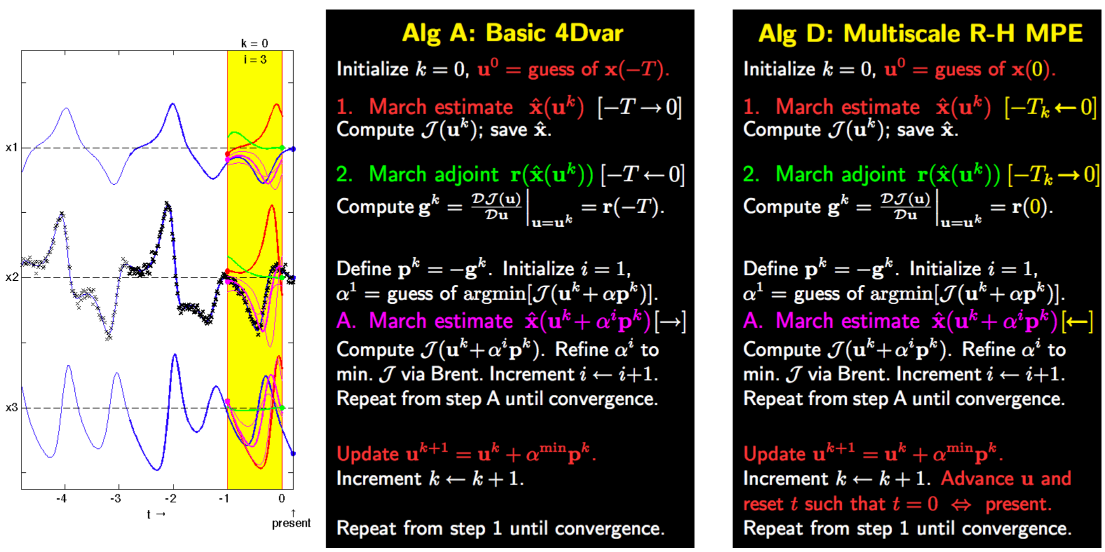

Figure 2: The basic 4Dvar algorithm for weather forecasting (center) and, after applying the three apparently minor but critical improvements discussed in this talk, the proposed multiscale receding-horizon model-predictive estimation approach (right). In the cartoon on the left, the blue curve corresponds to the underlying "truth" model, the black $\times$'s depict the noisy measurements of x_2, the yellow interval depicts the current optimization window, the red curve indicates the state estimate over this window, marched from $t = -T$ to $t = 0$ in step 1, the green cvurve indicates the adjoint marched the opposite direction over this window, from $t = 0$ to $t = -T$ in step 2, and the magenta curves indicate the state estimate for three trial values of α^i in step three. Note that time evolves as these computations are performed, as indicated by the gradual shifting of the yellow window to the left from the time marked as "present" as these computationally-intensive steps of the optimization proceed.