Variational Data Assimilation

by
Falk Feddersen
SIO

This lecture is a simple introduction to variational data assimilation using a mass spring oscillator (or a pendulum) as the system under study. The question this lecture addresses is how to make “optimal” estimates of the location of the mass-spring given both data and and dynamics. The mass-spring oscillator is an ideal system to study because (a) it is the *most* familiar dynamical system in existence, (b) analytic solutions can be written, and (c) it is linear.

Dynamical & Data Error

A perfect mass-spring will oscillate back and forth forever. However, no oscillator is perfect and some damping will occur whether due to a rusty spring or simply air friction. Therefore the dynamics have error and predictions based on a perfect mass spring oscillator dynamics will be in error. Now suppose one has *noisy* data of the position of the mass as it is oscillating back and forth. This lecture will demonstrate how to "optimally" combine the noisy data with the imperfect dynamics to yield "optimal" estimates of the location of the mass spring.

Simple mass-spring oscillator

Dynamics (Newton's 2nd Law)

$$x_{tt} + x = 0$$

- Called "Forward" model

Initial Conditions

At $t = 0$ the mass is released from rest at position $x = 1$

$$x(0) = 1$$

Initially at rest:

$$x_t(0) = 0$$

Solution

$$x(t) = \cos(t) \quad , \quad t \geq 0$$

- Called "Forward" solution.

Dynamics are **never** known perfectly. Allow for *errors* in the dynamcs. For example the spring is rusty.

$$x_{tt} + x = e(t)$$

$e(t)$ is errors in the dynamics.

Initial Conditions are also not known perfectly. For example the location of the release position $x(0)$, and the velocity $x_t(0)$ at the release point have a small error.

$$x(0) = 1 + f \quad , \quad x_t(0) = 0 + g$$

f and g are errors in I.C.

Suppose there are M noisy position data at various times,

$$d_i = x(t_i) + \epsilon_i,$$

where ϵ_i is the instrument error at time t_i.

Combine dynamics and data

Interested in position of mass for $0 \leq t \leq T$.

Minimize a combination of dynamical, initial condition, and data error,

$J = (\text{Dyn Error})^2 + (\text{I.C. Error})^2 + (\text{Data Error})^2$

$$J[x] = \int_0^T e^2(t)dt + f^2 + g^2 + \sum_{i=1}^{M} \epsilon_i^2$$

How to minimize?

With Calculus of Variations. Origins in Classical Mechanics (Principle of Least Action - Lagrangian Mechanics)

$$\delta J = J[x + \delta x] - J[x] = 0$$

Calculus of Variations

$$
\begin{aligned}
\delta J &= 2\int_0^T e(t)[(\delta x)_{tt} + \delta x]dt \\
&+ 2(x(0)-1)\delta x(0) + 2x_t(0)(\delta x(0))_t \\
&+ 2\sum_{i=1}^{M}\int_0^T (x(t)-d_i)\delta x \delta(t-t_i)dt
\end{aligned}
$$

Integrate by parts:

$$
\int_0^T e(t)(\delta x)_{tt}dt = [e(t)(\delta x)_t]_0^T - \int_0^T e_t(\delta x)_t dt =
$$

$$
= [e(t)(\delta x)_t]_0^T - [e_t(\delta x)]_0^T + \int_0^T e_{tt}\delta x dt
$$

$$
\begin{aligned}
\delta J &= \int_0^T \left[e_{tt} + e + \sum_{i=1}^{M}(x(t)-d_i)\delta(t-t_i)\right]\delta x\, dt \\
&+ e(T)(\delta x(T))_t + e_t(T)\delta x(T) \\
&+ (x(0)-(1+e_t(0)))\delta x(0) \\
&+ (x_t(0)-e(0))(\delta x(0))_t \\
&= 0
\end{aligned}
$$

Therefore to set δJ to zero:

$$e_{tt} + e + \sum_{i=1}^{M} (x(t) - d_i)\delta(t - t_i) = 0$$

$$e(T) = 0 \quad , \quad e_t(T) = 0$$

$$x_{tt} + x = e(t)$$

$$x(0) = 1 + e_t(0) \quad , \quad x_t(0) = e(0)$$

These are the Euler-Lagrange Equations

- ODE's, Linear, Coupled
- dynamical error $e(t)$ is the adjoint
- Solving this set gives the minimum of J
- Solutions are called "Inverse".

Note individual terms in cost function have different units. The relative influence of dynamics, IC, or data are not controlled. Solution: Weight each of the terms in the cost function.
Assume: all errors are Gaussian Random Variables with zero mean.

Dynamical Error Covariance

$$C_e(t,t') = \sigma_e^2 \exp\left[\frac{(t-t')^2}{T_e}\right]$$

σ_e^2 - zero lag dynamical error variance
T_e - decorrelation time scale (dynamical errors are correlated)

IC errors

σ_f^2 variance of mass position x at $t=0$
σ_g^2 variance of mass velocity x_t at $t=0$

Data Error

Identical & independent with variance σ_d^2

New Cost Function Weighted by (Co-) Variances

$$
\begin{aligned}
J[x] &= \int_0^T \int_0^T e(t) C_e^{-1}(t,t') e(t') dt' dt \\
&+ \sigma_f^{-2} f^2 + \sigma_g^{-2} g^2 + \sigma_d^{-2} \sum_{i=1}^{M} \epsilon_i^2
\end{aligned}
$$

- With Gaussian errors, minimizing $J[x]$ is equivalent to Maximum Likelihood Estimation.

- This is a *rational* quantity to minimize.

- Requires (prior) estimates of covariances!

- C_e requirement: symmetric and positive definite.

Minimizing New Cost Function

Define the adjoint λ as a convultion,

$$\lambda(t) = \int_0^T C_e^{-1}(t,t')e(t')dt'$$

where the inverse covariance $C_e{-1}$ is defined

$$\int_0^T C_e^{-1}(t,t'')C_e(t'',t')dt'' = \delta(t-t')$$

where δ is the Dirac delta function, analogous to matrix inverse $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.

New Euler-Lagrange Equations

$$\lambda_{tt} + \lambda + \sigma_d^{-2} \sum_{i=1}^{M} (x(t) - d_i)\delta(t - t_i) = 0$$

$$\lambda(T) = 0 \quad , \quad \lambda_t(T) = 0$$

$$x_{tt} + x = \int_0^T C_e(t,t')\lambda(t')dt'$$

$$x(0) = 1 + \sigma_f^2 \lambda_t(0) \quad , \quad x_t(0) = \sigma_g^2 \lambda(0)$$

What about Error Estimates?

Consider first no data.

Near minimum of $J[x]$

$$J[x+\delta x]-J[x]=\int_0^T\int_0^T e'(t)C_e^{-1}(t,t')e'(t')\,dtdt'+IC$$

where $e'(t)$ is dev. from min. Integrate by parts:

$$J[x+\delta x]-J[x]=\int_0^T\int_0^T \delta x C_x^{-1}(t,t')\delta x\,dtdt'$$

J is a infinitely dim. parabola (Quadratic Form).

$C_x(t,t')$ is the **prior** covariance of the errors in the mass location x, and is the inverse of the J curvature at the minimum.

$$C_x^{-1}(t,t')=\frac{d^4C_e^{-1}}{dt^2dt'^2}+\frac{d^2C_e^{-1}}{dt'^2}+\frac{d^2C_e^{-1}}{dt^2}+C_e^{-1}$$

for $0 < t, t' < T$.

- Prior C_x symmetric and positive definite.

Posterior Error Estimates

With data in the cost function, the *posterior* covariance $\widehat{C}_x$ becomes,

$$\widehat{C}_x^{-1}(t,t') = C_x^{-1}(t,t') + \sigma_d^{-2} \sum_{i=1}^{M} \delta(t-t_i)\delta(t'-t_i)$$

Since prior C_x positive definite, and
σ_d positive.

• Due to addition of data, the *posterior* covariance is **REDUCED**.

Example

True **damped** mass-spring oscillator ($\alpha = 0.015$):

$$x_{tt} + \alpha x_t + x = 0$$

- True solution: $x_{\text{true}} \approx \exp(-\alpha t)\cos[(1-\epsilon)t]$

$\rightarrow$ $\exp(-\alpha t)$ - Decaying envelope

$\rightarrow$ $\epsilon \ll 1$ - Tiny phase shift

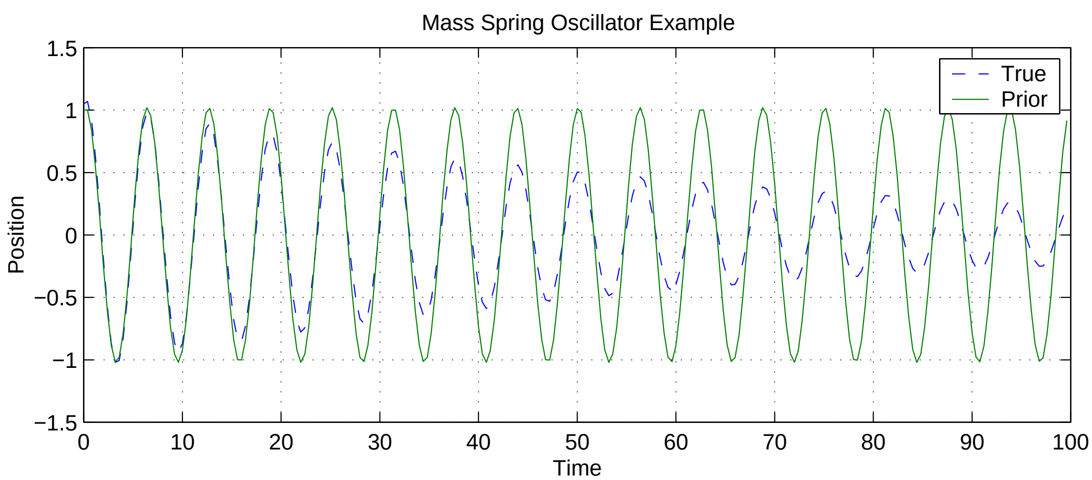
- Forward solution: $x_{\text{for}} = \cos(t)$

  Use Euler-Lagrange Equations to get

- Inverse solution: x_{inv}.

- Calculate prior & posterior x covariances.

True and Prior Solutions to Mass-Spring System

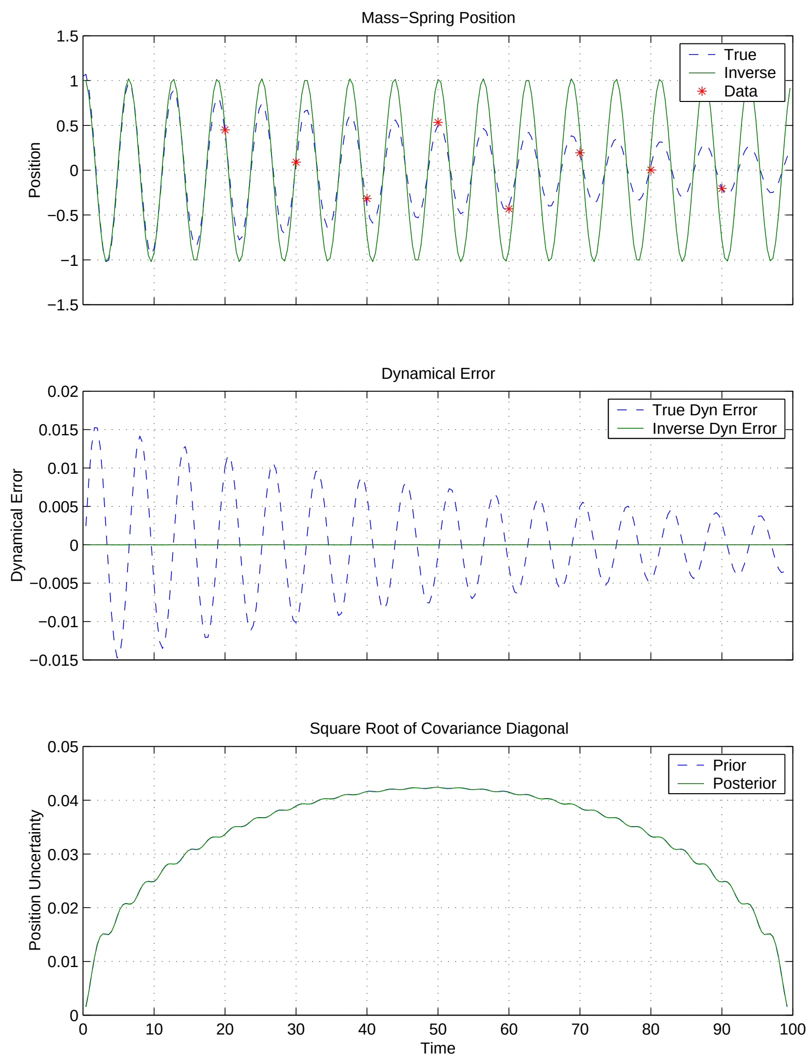


- True solution (dashed) is damped.

- Prior (Forward) solution (solid) is not.

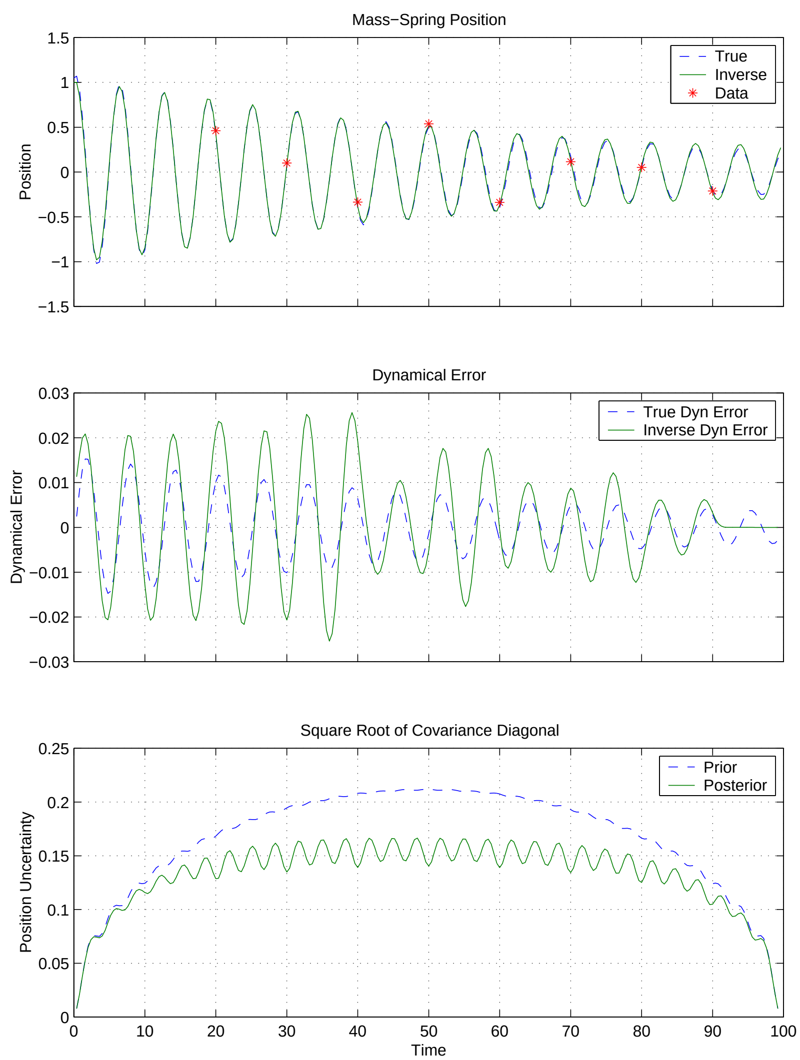
- Into this system we assimilate data.

Inverse Solution # 1.
Consider 8 Noisy Data
Assume Data Noise is **HUGE**: $\sigma_d \to \infty$



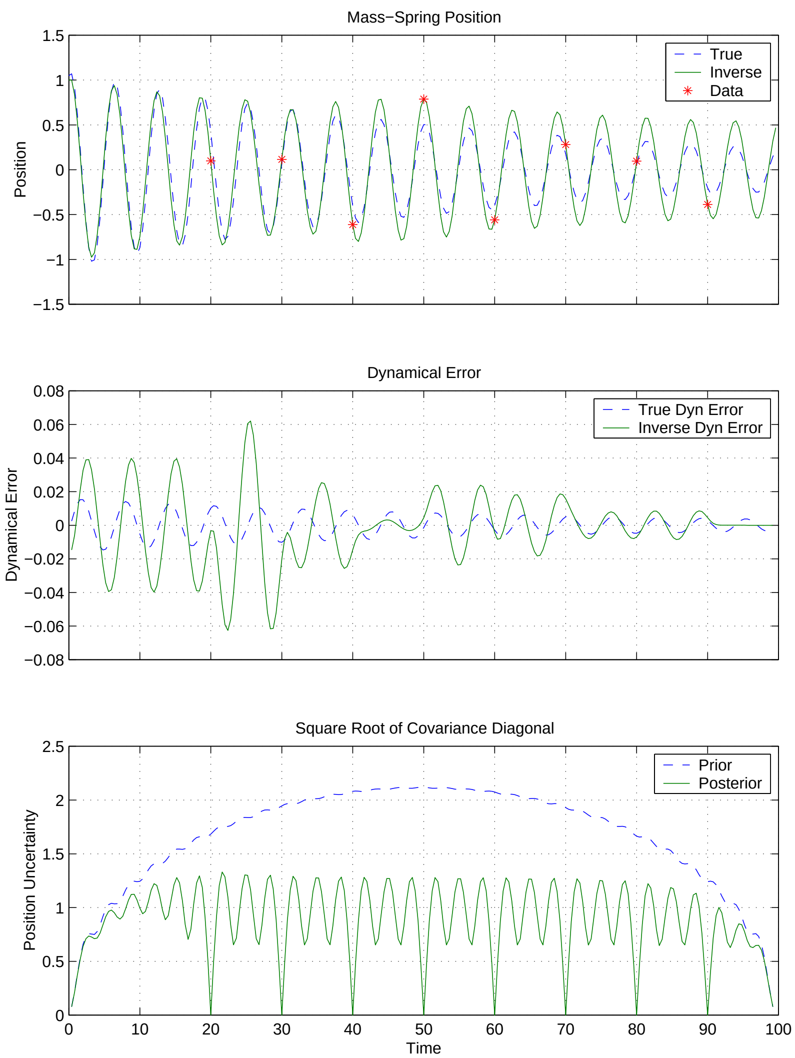
- Inverse solution equals prior solution.
- No reduction in uncertainty.

Inverse Solution # 2.
Medium Data Noise: $\sigma_d = 0.1$
Medium Dynamical Error: $\sigma_e = 0.05$



- Inverse solution **reproduces** true.
- Some reduction in uncertainty.

Inverse Solution # 3.
Assume Perfect Data: $\sigma_d = 0$
Data *actually* are noisy.



- Inverse solution overfits the noisy data.
- Spurious reduction in uncertainty.

Many many more issues

some are...

- How to solve Euler-Lagrange equations.
- How to accept/reject inverse solutions.
- Statistical interpretation of cost function minimization
- Nonlinear Systems.
- Parameter estimation.

Matlab programs that do the examples are available.

For more info contact `falk@coast.ucsd.edu`