Ocean Uncertainty Quantification Summer School

# Afternoon activity: build a Kalman filter

Aneesh Subramanian, Sarah Williamson

July 18, 2024

# Introduction

- For the afternoon activity we'll be writing code that implements a Kalman filter

# Introduction

- For the afternoon activity we'll be writing code that implements a Kalman filter
- The foundation comes from "Potential artifacts in conservation laws and invariants inferred from sequential state estimation" by Wunsch et al. (2023)

# Introduction

- For the afternoon activity we'll be writing code that implements a Kalman filter
- The foundation comes from "Potential artifacts in conservation laws and invariants inferred from sequential state estimation" by Wunsch et al. (2023)
- Goal is to adjust data, uncertainties, known/unknown assumptions etc. to see the impact on state reconstruction

# Notation summary

Let $0 \leq t \leq t_f$ denote model time where $t_f = N_t \Delta t$ for timestep $\Delta t$. Define

- $\boldsymbol{x}(t)$: state vector
- $\boldsymbol{A}$: state transition matrix
- $\boldsymbol{q}(t)$: perturbation/disturbance vector
- $\boldsymbol{B}$: disturbance distribution matrix
- $\boldsymbol{u}(t)$: control terms (unknown forcing)
- $\boldsymbol{\Gamma}$: control terms distribution matrix

Combining the above we define a time-evolving system,

$$\boldsymbol{x}(t + \Delta t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{q}(t) + \boldsymbol{\Gamma}\boldsymbol{u}(t)$$

# Notation summary

Let $0 \leq t \leq t_f$ denote model time where $t_f = N_t \Delta t$ for timestep $\Delta t$. Define

- $\boldsymbol{x}(t)$: state vector
- $\boldsymbol{A}$: state transition matrix
- $\boldsymbol{q}(t)$: perturbation/disturbance vector
- $\boldsymbol{B}$: disturbance distribution matrix
- $\boldsymbol{u}(t)$: control terms (unknown forcing)
- $\boldsymbol{\Gamma}$: control terms distribution matrix

Combining the above we define a time-evolving system,

$$\boldsymbol{x}(t + \Delta t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{q}(t) + \boldsymbol{\Gamma}\boldsymbol{u}(t)$$

# Model data

Define

- $\boldsymbol{y}(t)$: data vector
- $\boldsymbol{n}(t)$: noise vector

Combining the above gives form for observations/data,

$$\boldsymbol{y}(t) = \boldsymbol{E}\boldsymbol{x}(t) + \boldsymbol{n}(t)$$

where $\boldsymbol{E}$ distributes entries of $\boldsymbol{x}$

# Kalman filter equations

$$\tilde{\boldsymbol{x}}(t,-) = \boldsymbol{A}\tilde{\boldsymbol{x}}(t-\Delta t) + \boldsymbol{B}\boldsymbol{q}(t-\Delta t)$$

# Kalman filter equations

$$\tilde{\boldsymbol{x}}(t,-) = \boldsymbol{A}\tilde{\boldsymbol{x}}(t-\Delta t) + \boldsymbol{B}\boldsymbol{q}(t-\Delta t)$$
$$\boldsymbol{P}(t,-) = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^T + \Gamma \boldsymbol{Q} \Gamma^T$$

# Kalman filter equations

$$\tilde{\boldsymbol{x}}(t,-) = \boldsymbol{A}\tilde{\boldsymbol{x}}(t-\Delta t) + \boldsymbol{B}\boldsymbol{q}(t-\Delta t)$$
$$\boldsymbol{P}(t,-) = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^T + \Gamma\boldsymbol{Q}\Gamma^T$$
$$\boldsymbol{K}(t) = \boldsymbol{P}(t,-)\boldsymbol{E}^T\left(\boldsymbol{E}\boldsymbol{P}(t,-)\boldsymbol{E}^T + \boldsymbol{R}\right)^{-1}$$

# Kalman filter equations

$$
\begin{aligned}
\tilde{\boldsymbol{x}}(t,-) &= \boldsymbol{A}\tilde{\boldsymbol{x}}(t-\Delta t) + \boldsymbol{B}\boldsymbol{q}(t-\Delta t) \\
\boldsymbol{P}(t,-) &= \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^T + \Gamma\boldsymbol{Q}\Gamma^T \\
\boldsymbol{K}(t) &= \boldsymbol{P}(t,-)\boldsymbol{E}^T\left(\boldsymbol{E}\boldsymbol{P}(t,-)\boldsymbol{E}^T + \boldsymbol{R}\right)^{-1} \\
\tilde{\boldsymbol{x}}(t) &= \tilde{\boldsymbol{x}}(t,-) + \boldsymbol{K}(t)\left(\boldsymbol{y}(t) - \boldsymbol{E}\tilde{\boldsymbol{x}}(t,-)\right)
\end{aligned}
$$

# Kalman filter equations

$$
\begin{aligned}
\tilde{\boldsymbol{x}}(t,-) &= \boldsymbol{A}\tilde{\boldsymbol{x}}(t-\Delta t) + \boldsymbol{B}\boldsymbol{q}(t-\Delta t) \\
\boldsymbol{P}(t,-) &= \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^T + \Gamma \boldsymbol{Q} \Gamma^T \\
\boldsymbol{K}(t) &= \boldsymbol{P}(t,-)\boldsymbol{E}^T \left(\boldsymbol{E}\boldsymbol{P}(t,-)\boldsymbol{E}^T + \boldsymbol{R}\right)^{-1} \\
\tilde{\boldsymbol{x}}(t) &= \tilde{\boldsymbol{x}}(t,-) + \boldsymbol{K}(t)\left(\boldsymbol{y}(t) - \boldsymbol{E}\tilde{\boldsymbol{x}}(t,-)\right) \\
\boldsymbol{P}(t) &= \boldsymbol{P}(t,-) - \boldsymbol{K}(t)\boldsymbol{E}\boldsymbol{P}(t,-)
\end{aligned}
$$

# Kalman filter equations

$$\tilde{\boldsymbol{x}}(t,-) = \boldsymbol{A}\tilde{\boldsymbol{x}}(t-\Delta t) + \boldsymbol{B}\boldsymbol{q}(t-\Delta t)$$
$$\boldsymbol{P}(t,-) = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^T + \Gamma\boldsymbol{Q}\Gamma^T$$
$$\boldsymbol{K}(t) = \boldsymbol{P}(t,-)\boldsymbol{E}^T\left(\boldsymbol{E}\boldsymbol{P}(t,-)\boldsymbol{E}^T + \boldsymbol{R}\right)^{-1}$$
$$\tilde{\boldsymbol{x}}(t) = \tilde{\boldsymbol{x}}(t,-) + \boldsymbol{K}(t)\left(\boldsymbol{y}(t) - \boldsymbol{E}\tilde{\boldsymbol{x}}(t,-)\right)$$
$$\boldsymbol{P}(t) = \boldsymbol{P}(t,-) - \boldsymbol{K}(t)\boldsymbol{E}\boldsymbol{P}(t,-)$$

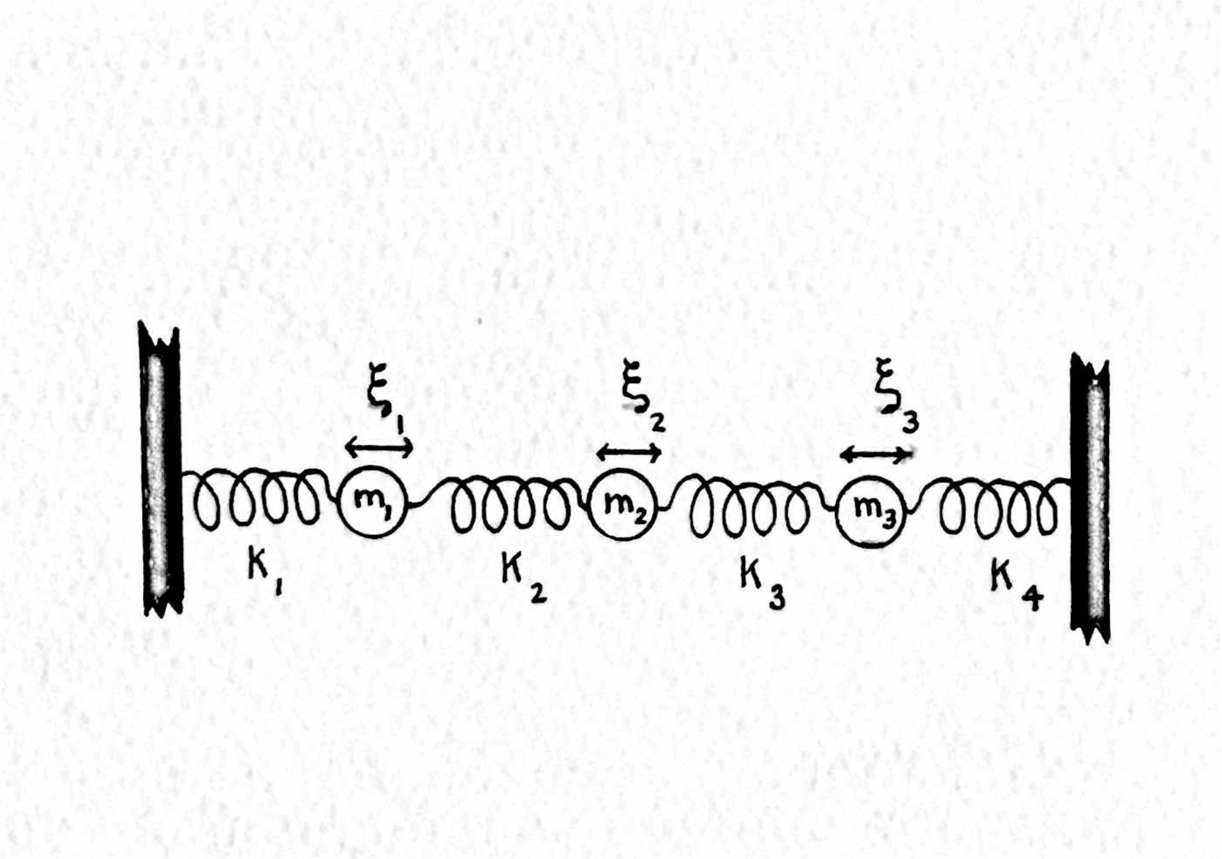
**Known (deterministic) forcing term**

# Kalman filter equations

$$\tilde{\boldsymbol{x}}(t,-) = \boldsymbol{A}\tilde{\boldsymbol{x}}(t-\Delta t) + \boldsymbol{B}\boldsymbol{q}(t-\Delta t)$$
$$\boldsymbol{P}(t,-) = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^T + \boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^T$$
$$\boldsymbol{K}(t) = \boldsymbol{P}(t,-)\boldsymbol{E}^T\left(\boldsymbol{E}\boldsymbol{P}(t,-)\boldsymbol{E}^T + \boldsymbol{R}\right)^{-1}$$
$$\tilde{\boldsymbol{x}}(t) = \tilde{\boldsymbol{x}}(t,-) + \boldsymbol{K}(t)\left(\boldsymbol{y}(t) - \boldsymbol{E}\tilde{\boldsymbol{x}}(t,-)\right)$$
$$\boldsymbol{P}(t) = \boldsymbol{P}(t,-) - \boldsymbol{K}(t)\boldsymbol{E}\boldsymbol{P}(t,-)$$

**Unknown forcing term**

# Three-mass spring oscillator



# Equations of motion

$$m\frac{d^2\xi_1}{dt^2} + k\xi_1 + k(\xi_1 - \xi_2) + r\frac{d\xi_1}{dt} = q_1(t)$$
$$m\frac{d^2\xi_2}{dt^2} + k\xi_2 + k(\xi_2 - \xi_1) + k(\xi_2 - \xi_3) + r\frac{d\xi_2}{dt} = q_2(t)$$
$$m\frac{d^2\xi_3}{dt^2} + k\xi_3 + k(\xi_3 - \xi_2) + r\frac{d\xi_3}{dt} = q_3(t)$$

# Re-writing in the form $x(t) = Ax(t - \Delta t) + Bq(t)$

- Step 1: Reduce the system of equations to a first order system

# Re-writing in the form $x(t) = Ax(t - \Delta t) + Bq(t)$

- Step 1: Reduce the system of equations to a first order system
- Step 2: Discretize the system, we recommend Euler's method for the time discretization

# Re-writing in the form $x(t) = Ax(t - \Delta t) + Bq(t)$

- Step 1: Reduce the system of equations to a first order system
- Step 2: Discretize the system, we recommend Euler's method for the time discretization

Hint for 1: Introduce the vector

# Re-writing in the form $x(t) = Ax(t - \Delta t) + Bq(t)$

- Step 1: Reduce the system of equations to a first order system
- Step 2: Discretize the system, we recommend Euler's method for the time discretization

Hint for 1: Introduce the vector

$$\boldsymbol{x}(t) = \begin{pmatrix} \xi_1(t) \\ \xi_2(t) \\ \xi_3(t) \\ d\xi_1/dt \\ d\xi_2/dt \\ d\xi_3/dt \end{pmatrix}$$

# Re-writing in the form $x(t) = Ax(t - \Delta t) + Bq(t)$

Let

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{I}_3 & \Delta t \boldsymbol{I}_3 \\ \Delta t \boldsymbol{K}_c & \boldsymbol{I}_3 + \Delta t \boldsymbol{R}_c \end{pmatrix}$$

# Re-writing in the form $x(t) = Ax(t - \Delta t) + Bq(t)$

Let

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{I}_3 & \Delta t \boldsymbol{I}_3 \\ \Delta t \boldsymbol{K}_c & \boldsymbol{I}_3 + \Delta t \boldsymbol{R}_c \end{pmatrix}$$

We chose the forcing

$$\boldsymbol{q}(t) = q_1(t) = 0.1 \cos\left(\frac{2\pi t}{2.5r}\right) + \varepsilon(t)$$

where $\varepsilon(t)$ is white noise with variance $0.1^2$

# Re-writing in the form $x(t) = Ax(t - \Delta t) + Bq(t)$

Let

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{I}_3 & \Delta t \boldsymbol{I}_3 \\ \Delta t \boldsymbol{K}_c & \boldsymbol{I}_3 + \Delta t \boldsymbol{R}_c \end{pmatrix}$$

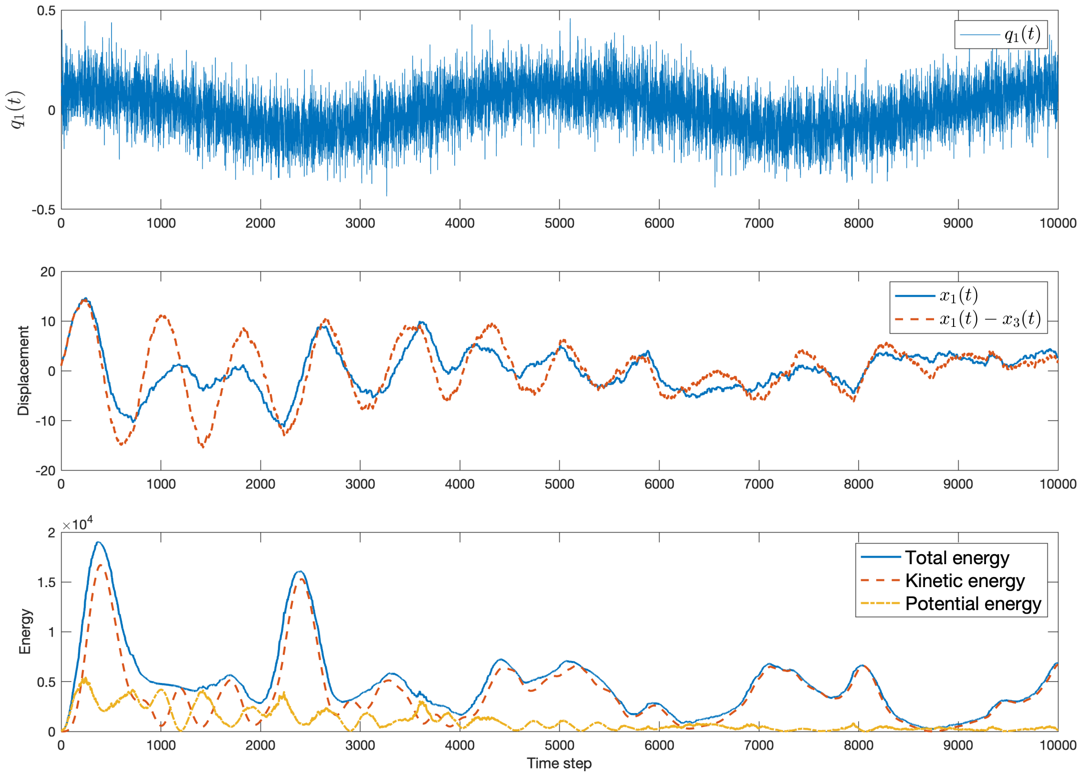
We chose the forcing

$$\boldsymbol{q}(t) = q_1(t) = 0.1 \cos\left(\frac{2\pi t}{2.5 r}\right) + \varepsilon(t)$$

where $\varepsilon(t)$ is white noise with variance $0.1^2$
This gives the final timestepping form

$$\boldsymbol{x}(t + \Delta t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{q}(t)$$

# Forced solution



# Ideas for reconstruction:

- Elements of the state vector

# Ideas for reconstruction:

- Elements of the state vector
- Energy $\mathcal{E}(t)$ (a diagnostic quantity):

$$\mathcal{E}(t) = \frac{1}{2}\left\{ \begin{pmatrix} x_4(t) \\ x_5(t) \\ x_6(t) \end{pmatrix}^T \begin{pmatrix} x_4(t) \\ x_5(t) \\ x_6(t) \end{pmatrix} - \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix}^T \boldsymbol{K}_c \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} \right\}$$

# Sources of uncertainty across all experiments

- Noisy data (i.e. $\boldsymbol{y}(t) = \boldsymbol{x}(t) + \boldsymbol{n}(t)$)

# Sources of uncertainty across all experiments

- Noisy data (i.e. $\boldsymbol{y}(t) = \boldsymbol{x}(t) + \boldsymbol{n}(t)$)
- Imperfect forcing $0.5\boldsymbol{q}(t) - \varepsilon(t)$, i.e.:

$$\boldsymbol{q}(t) = 0.1 \cos\left(\frac{2\pi t}{2.5r}\right) + \varepsilon(t)$$

This value will be half the true value

# Sources of uncertainty across all experiments

- Noisy data (i.e. $\boldsymbol{y}(t) = \boldsymbol{x}(t) + \boldsymbol{n}(t)$)
- Imperfect forcing $0.5\boldsymbol{q}(t) - \varepsilon(t)$, i.e.:

$$\boldsymbol{q}(t) = 0.1 \cos\left(\frac{2\pi t}{2.5r}\right) + \varepsilon(t)$$

Stochastic part will be fully unknown

# Experiment ideas

- Accurate observations of all elements of $\boldsymbol{x}$

# Experiment ideas

- Accurate observations of all elements of $\boldsymbol{x}$
- Fixed position, $x_3(t) = x_3 = 2$

# Experiment ideas

- Accurate observations of all elements of $\boldsymbol{x}$
- Fixed position, $x_3(t) = x_3 = 2$
- Observations of averages