Data Assimilation and Climate
AIMS Rwanda, March, 2020

Lab Activities
Introduction to Calculus of Variation

Consider a physical process with state variables x_i generated according to the rule $x_{i+1} = f(x_i)$ where $f(x) = 4x(1-x)$. The data model is $y_i = x_i + \epsilon_i$ where the ϵ_i are independent random variables with a $N(0, \sigma^2)$ distribution.

1. Start with $x_1 = 0.2$ and form the state variables x_i, for $i = 2, 3, 4$.
2. Form the data y_i, $i = 2, 3$, with $\sigma = 0.001$. In order to do this you will need ϵ_i, random variables from a normal distribution with mean 0 and standard deviation σ. You can get ϵ_i with the Python function
np.random.normal(mean,standard deviation,N) ,
where N is the size of the array (in our case $N = 4$). Plot the four data points as points (not lines), e.g. in Python
plot(y,'b',markersize=4)* .
3. Construct the cost function $\mathcal{J}(x_2; y_2, y_3) = (x_2 - y_2)^2 + (x_3 - y_3)^2$ so that it does not depend on x_3. In other words, replace x_3 with $4x_2(1 - x_2)$. Given the data you found in 2., and for $x_2 \in [-0.1, 1.1]$, plot $\mathcal{J}$ vs x_2. Do you anticipate any difficulties in finding the minimum of $\mathcal{J}$? Please post your thoughts on the activities channel in Slack.
4. Now let's assume that the standard deviation in the data is much larger, say $\sigma = 0.9$. Form new data y_i and cost function $\mathcal{J}(x_2; y_2, y_3)$. Plot the new cost function together with the old cost function. Describe your observations., in particular will the minimum occur at different places when the data errors change?

Something to think about: If you had more data and more states, say $i = 2, \ldots, 6$, how would you form the cost function? What would the cost function look for any general function $x_{i+1} = f(x_i)$?