1 HW2 Hint

1.1 Problem 1

1.1.1 Part 2

It is helpful to realize that x and y depend not only on t but on the four parameters $p = (\alpha, \beta, \gamma, \delta)$.

Thus it is reasonable to evolve not only u=[x, y] with time but also the eight variables in the 2x4 matrix:

$$\frac{\partial u}{\partial p} = \begin{pmatrix} \frac{\partial x}{\partial \alpha} & \frac{\partial x}{\partial \beta} & \frac{\partial x}{\partial \gamma} & \frac{\partial x}{\partial \delta} \\ \frac{\partial y}{\partial \alpha} & \frac{\partial y}{\partial \beta} & \frac{\partial y}{\partial \gamma} & \frac{\partial y}{\partial \delta} \end{pmatrix}.$$

Thus we are evolving 10 variables in total. I'm wondering if it matters if we start these eight variables at 0 at t=0 or not?

Here

$$f(u, p, t) = \begin{pmatrix} \alpha x - \beta xy \\ -\gamma y + \delta xy \end{pmatrix}.$$

You will need the Jacobian of f with respect to x and y:

$$\frac{\partial f}{\partial u} = \begin{pmatrix} \alpha - \beta y & -\beta x \\ \delta y & -\gamma + \delta x \end{pmatrix},$$

and also the Jacobian of f with respect to $\alpha, \beta, \gamma, \delta$:

$$\frac{\partial f}{\partial p} = \begin{pmatrix} x & -xy & 0 & 0 \\ 0 & 0 & -y & xy \end{pmatrix}.$$

Note that the resulting system does not have a nice analytical solution since x and y are functions of t. Instead, use your integrator from part 1 for solving the new combined system.

1.1.2 Part 3

First you will need to write down the loss function you want to minimize. You are asked to use the L2-norm of the difference between your computed solution $u(t_i)$ and the original solution from part 1 $\hat{u}(t_i)$ you are trying to recreate (the training data if you will). The loss function $L(u)$ then looks as follows:

$$L(u) = \sum_i (u(t_i) - \hat{u}(t_i))^2$$

You want to minimize this function via gradient descent, so you need to find the gradient w.r.t. the parameters p (α, β, γ, and δ). Use the chain rule:

$$\frac{\partial L}{\partial p} = \sum_i \frac{\partial L}{\partial u(t_i)} \cdot \frac{\partial u(t_i)}{\partial p}$$

$\frac{\partial L}{\partial u(t_i)}$ is straightforward to derive from the previous equation and $\frac{\partial u(t_i)}{\partial p}$ is exactly what you were supposed to find a way to calculate numerically in part 2.