Uncertainty Quantification course
Homework assignment 7

November 2, 2023

We discuss this exercise in the class meeting on 8 November 2023.

We have a very simple epidemiological model, a so-called SEIR model (for Susceptible, Exposed, Infected, Recovered). It consists of 4 coupled ODEs for $S(t), E(t), I(t)$ and $R(t)$, the number of individuals in each of the four categories (also called "compartments"). The total population consists of $N = S(t) + E(t) + I(t) + R(t)$ individuals; we assume that N is constant in time. In this assignment you use stochastic collocation to explore the uncertainty in the model output due to uncertainty in **three** input parameters: the reproduction number R_0, the incubation period T and the infectious period τ.

On Canvas you will find two Matlab codes `SEIRmodel.m` and `SEIRscript.m` for numerical integration of a SEIR model. As mentioned, the model has three parameters: the reproduction number R_0, the incubation period T (in days) and the infectious period τ (also in days). The output consists of timeseries for S, E, I and R. The population size is set at $N = 10^5$, the initial condition is set at $E(0) = 0$, $I(0) = 10$, $R(0) = 0$ and $S(0) = N - E(0) - I(0) - R(0)$. The code is based on a model from https://cs.uwaterloo.ca/∼paforsyt/SEIR.html. The time integration runs from $t = 0$ to $t = 500$ with timestep $dt = 0.01$.

We assume R_0 has uniform distribution on the interval $[1.5, 3.0]$. Furthermore, T has uniform distribution on the interval $[5, 10]$, and τ has a beta distribution on the interval $[1, 14]$, with probability density function $\rho(\tau) \propto (\tau - 1)(14 - \tau)$.

Let Q be the Quantity of Interest, defined as the maximum of $I(t)$ over time, i.e. $Q = \max_t I(t)$. Thus, Q is the peak number of infected individuals (see also the code `SEIRscript.m`). It is dependent on the parameters R_0, T, τ.

Use stochastic collocation to construct an approximation $\tilde{Q}(R_0, T, \tau)$ of the exact $Q(R_0, T, \tau)$. To construct the approximation, don't use a built-in / library code for interpolation, instead write your own code using Lagrange interpolating functions. For the nodes, you can use the Clenshaw-Curtis nodes. Furthermore, you can use full tensor grids for your approximation, you don't have to go to sparse grids.

Once you have built your approximation $\tilde{Q}(R_0, T, \tau)$, use Monte Carlo sampling to assess the probability distribution of $\tilde{Q}$. Make a histogram plot to show the distribution of the MC samples. What are its mean and standard deviation? Can you give an estimate of the probability that $\tilde{Q}$ will be higher than 15000?