PRINCIPLES OF COMPLEX SYSTEMS

ASSIGNMENT 7

January 6, 2023

Maxfield Green
University of Vermont
Comp. Systems and Data Sci.

0.1 Problem 1

Determine $p >= (x)$ and $p >= (A)$ for the following :

(a) $p(x) = cx^{-(q+1)}$

$$Cx^{-(q+1)} = CA^{-\gamma} \tag{1}$$
$$x = A^{\frac{\gamma}{q+1}} \tag{2}$$
$$\tag{3}$$
$$P \geq (A) = \int_{A^{\frac{\gamma}{(q+1)}}}^{\infty} x^{-(q+1)} dx \tag{4}$$
$$P \geq (x) = x^{-q} \tag{5}$$
$$P \geq (A) = C\frac{A^{\gamma(1+\frac{1}{q})}}{q} \tag{6}$$

(b) $p(x) = ce^{-x}$

$$Ce^{-x} = CA^{-\gamma} \tag{7}$$
$$x = -\ln A^{-\gamma} \tag{8}$$
$$\tag{9}$$
$$P \geq (A) = \int_{\ln A^{-\gamma}}^{\infty} e^{-x} dx \tag{10}$$
$$P \geq (A) = 0 - (-e^{ln(A^{-\gamma})}) = CA^{-\gamma} \tag{11}$$
$$P \geq (X) = e^{-x} \tag{12}$$
$$\tag{13}$$

(c) $p(x) = ce^{-x^2}$

This one is very sneaky. Nice try, slipping in the Gaussian Integral. Because the function technically cannot be integrated definitely. There is some talk of polar coordinates on Wikipedia.

However, I suggest an alternate approach.

$$CA^{-\gamma} = Ce^{-x^2} \tag{14}$$

$$X = -ln(A^{\gamma})^{1/2} \tag{15}$$

$$\tag{16}$$

We have our lower bound, X. Integrating by parts gives us: (17)

$$uv = x^{-1}e^{x^2}, u = x^{-1} v = e^{-x^2} \tag{18}$$

$$\tag{19}$$

$$\int_x^{\infty} x^{-1}x2e^{-x^2}dx = x^{-1}e^{-x^2} - \int_x^{\infty} -\frac{e^{-x^2}}{x^2}dx \tag{20}$$

$$\int_x^{\infty} e^{-x^2}dx = e^{-x^2}x^{-1}\frac{1}{2} + \int_x^{\infty} \frac{e^{-x^2}}{x^2}dx \tag{21}$$

The third term is effectively zero under the bounds of integration when we are only looking at the tail, as specified in the question.
Evaluating our limits of integration gives :

$$\int_x^{\infty} e^{-x^2}dx = -2e^{-x^2}|_{(-ln(A^{-\gamma})^{-1/2}}^{\infty} \tag{22}$$

$$P \geq (A) = \int_x^{\infty} e^{-x^2}dx = CA^{-\gamma}ln(A)^{-1/2} \tag{23}$$

$$\tag{24}$$

0.2 Problem 2

The Discrete version of HOT theory.
Given that:

Cost : Expected size of fire in a 2-D lattice

$$C_{fire} \propto \sum_{i=1}^{Nsites} p_i a_i \tag{25}$$

The constraint for building and maintaining (d-1) dimensional firewalls in d-dimensions in is:

$$C_{firewalls} \propto \sum_{i=1}^{Nsites} a_i^{(d-1)/d} a_i^- 1 \tag{26}$$

This problem is perfectly suited for the Lagrangian Multiplier technique. We are looking to maximize the yield of the forest with the constraint of our finite resource being fire breaks. The setup for this looks like:

$$\nabla C_{fire} = \lambda \nabla C_{firewalls} \tag{27}$$

Taking the gradients we get:

$$p_i = \lambda a^{-(d+1)/d} \tag{28}$$

$$p_i = \lambda a^{-(1+1/d)} \tag{29}$$

$$p_i \propto a^{-(1+1/d)} \tag{30}$$

0.3 Problem 3

Discrete Highly Optimized Tolerance in 2 Dimensions. Timber yield is optimized in the presence of known fire risk. We find timber yield curves as a function of information level the model is given. We illustrate a system that is produces high yields that are robust to "designed for" uncertainty but highly sensitive to unexpected changes and power law distributions, a characteristic of "complex systems"!

In this model, successful strategies are rewarded to minimize the cost of timber loss in the event of a fire. By optimizing yield of timber, at certain forest densities, barriers that minimize the size of a fires are created. Barriers will isolate areas of high spark probability, leaving large connected components at risk for massive but unlikely fire events.

We start with a given probability space for fire causing emissions over a 2D grid of size NxN. The spacial distribution of probabilities of a spark emission at location (i, j) is given by:

$$P(i, j) = e^{-i/l} e^{-j/l} \tag{31}$$

where

$$l = \frac{\sum_{i=1}^{N}}{10} \tag{32}$$

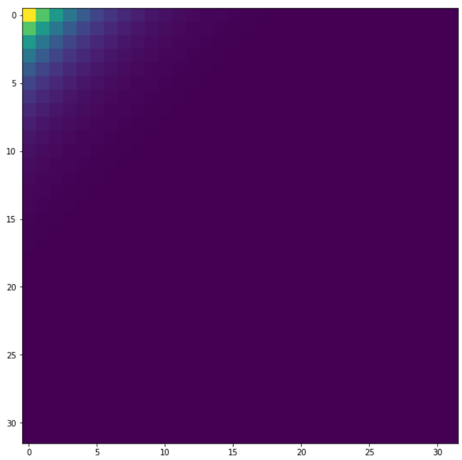
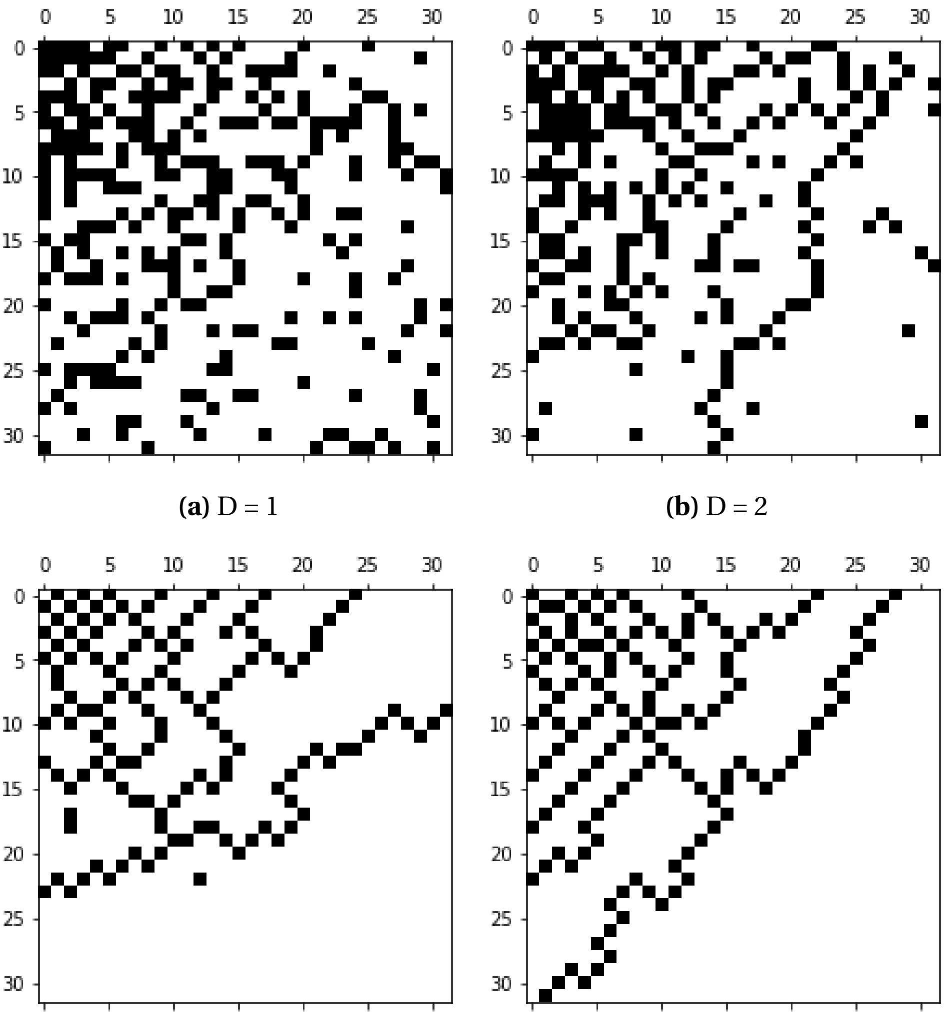


Figure 1: Probability of Spark over landscape. Light colors indicate high spark probability.

The forest is initiated void of trees. Given discrete time, the density of the forest increases by one tree at a time. D sites are tested for tree growth prior to final planting. The tree location. The tree that is planted produces the maximum yield of un-burned trees in the event of a fire is chosen. Tree yield is defined as the proportion of trees remaining in forest after a fire. When a spark hits a forest site that contains a tree, the fire spreads to every tree neighboring site in the connected component. As the density of the forest rises a maximum timber yield is reached. At this point, any trees added to the forest will connect two separate components, increasing the cost of a potential fire event. Our system optimizes the economic gain of timber maximization with the cost of large connect components of trees that could burn down the forest with a single spark. Below, the forest at peak yield on a 32x32 grid are plotted for varying values of D.

////



(a) D = 1

(b) D = 2

(c) D = L

(d) D = L^2

Figure 2: Forests at peak yield for varying values of D.

A key piece here is that the algorithm does not ever get a full look at the probability distribution space, just little glimpses at a time. It picks the best outcome given the information that is available. The information levels vary with D the number of potential tree plants tested. In the case (d) when D = L^2, the forest gets close to being greedily optimized. It has the possibility of seeing every possible combination. However, there is still an element of randomness in the model. It is possible that one of those peaks at a potential future will be the same, thus denying the algorithm from seeing all potential futures. In this way, the $D = L^2$ case has the potential to be deterministic with some probability.

The yield curves for these configurations are displayed below.

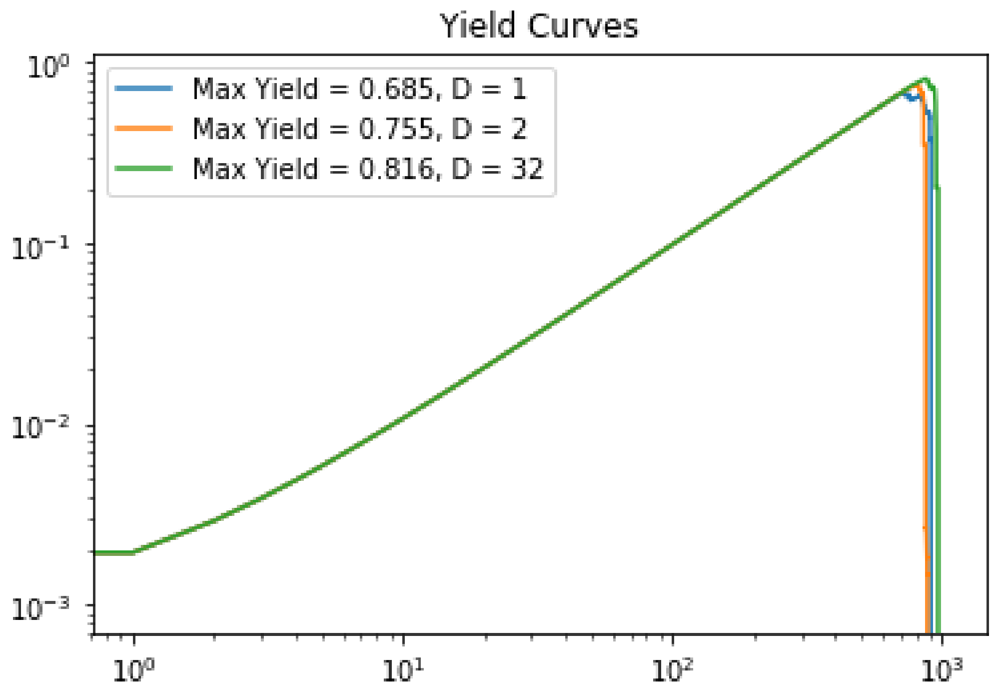


Figure 3: Yield curves for varying design parameter D.

We see that the highest yield results from when D = 32. This makes a lot of sense, when the model has more information about the risk landscape, it can optimize much better. I would however expect a larger difference between D = 1,2 and D = 32.