

```
In [4]: %pylab inline
plt.xkcd()
```

Populating the interactive namespace from numpy and matplotlib

```
Out[4]: <matplotlib.rc_context at 0x1074375d0>
```

TB in Cattle and Badgers

Overview of Prior Work

From studying an epidemic in bovine tuberculosis (TB) in 2001, D.R. Cox *et al.* came up with a simple model for tracking the number of infected cattle as it relates to the surrounding population of badgers. Their derivation took the following steps:

Let $C(t)$ (the number of infected cattle per km^2) and $B(t)$ (the number of infected badgers per km^2) be the rows of a vector \vec{X} . Also, let a_{ij} be the rate (per year, typically) that an animal of species j infects an animal of species i . Furthermore, the removal of animals of species i is governed by some equation $r_i(t)$ and the restocking of that species is governed by $s_i(t)$. Finally, letting $\vec{A} = a_{ij}$, $\vec{R} = \text{diag}(r_1, r_2)$, and $\vec{S} = [s_1, s_2]^T$ yields:

$$\dot{\vec{X}} = (\vec{A} - \vec{R})\vec{X} + \vec{S}$$

which can be solved:

$$\vec{X} = \vec{B}^{-1}\vec{S} + e^{\vec{B}t}(\vec{X}_0 - \vec{B}^{-1}\vec{S})$$

where $\vec{B} = \vec{R} - \vec{A}$ and $\vec{X}_0 = \vec{X}(0)$.

The matrix exponent can be found by:

$$\vec{X} = \vec{B}^{-1}\vec{S} + \vec{H}\text{diag}(e^{-\lambda_1 t}, e^{-\lambda_2 t})(\vec{X}_0 - \vec{B}^{-1}\vec{S})$$

where $\vec{B} = \vec{H}\vec{\Lambda}\vec{H}^{-1}$ (the spectral decomposition is used in order to find the matrix exponential).

This, however, assumes that there are no external factors on the number of infected badgers. If there were (culling would be one such example) the equation for the derivative with respect to time for the number of infected cattle is:

$$\dot{C} = (a_{11} - r_1)C + a_{12}B + s_1$$

There are several parameters that inform the behavior of these functions. If routine testing happens every d years and the sensitivity of the skin test for TB is p , then the average interval between incidence and detection is

$$\mu_d = [1/2 + (1 - p)/p]d.$$

This equation, however, assumes a very simple geometric distribution over the number of retests which likely overestimates μ_d .

A herd in which a reactor is found (that is, one in which a cow that got a positive result from the TB test) is tested again. To be sure, this means that as the sensitivity increases, the amount of time it takes to test increases as well (as the increased number of positive results will thereby lead to more testing). The effect that the sensitivity of the test has on the effecting testing interval is given by the following ratio:

$$\mu_d/(d/2) = 1 + 2(1 - p)/p.$$

The average amount of time that a herd is in the absense of a detected breakdown, μ_n , is based one the probability density function, $f(x)$, and its cumulative distribution function $F(x)$ on the distribution of life in a disease-free herd. The mean and variance of $f(x)$ are μ_f and σ_f^2 respectively. According to D.R. Cox *et al.* in a 1965 paper on stochastic proceses, the PDF of time between infection and routine removal is just $[1 - F(x)]/\mu_f$. The expected value of which is somehow $\mu_f(1 + \sigma_f^2)/2$. Finally, this informs the removal rate for infected cattle:

$$r_1 = \mu_n^{-1} + \mu_d^{-1}$$

Thus, the removal rate for infected cattle is related both to the routine removal of cattle that are infected but not known to be as well as the time it takes to identify infected cattle.

Looking at equilibrium points $\vec{X}^* = (C^*, B^*)$ reveals that

$$C^* = \frac{s_1 + a_{12}B^*}{r_1 - a_{11}}$$

which means that there can be no equilibrium unless the removal rate of infected cattle, r_1 is greater than the rate at which cattle infect eachother, a_{11} . If badgers are then culled, however, to sustain a badger population of $(1 - k)B^*$, then the equilibrium point for the cattle is:

$$C^k = \frac{s_1 + (1 - k)a_{12}B^*}{r_1 - a_{11}}.$$

From this, we can estimate the how much of an effect badger culling has on the number of infected cattle:

$$C^k/C^* = \frac{1 + (1 - k)I_b}{1 + I_b}$$

where $I_b = a_{12}B^*/s_1$, or the "ratio of the number of cattle infected by badgers per year to the number of infected cattle brought in from outside of the system per year." This value, unfortunatly, is not known (unlike a lot of these other values for which there is actual data).

At this point in the paper, D.R. Cox *et al.* go into the time time-dependent solutions. In order to do this, they solve the differential equaton above, yeilding:

$$C(t) = C_0 + a_{12}B_0k(\lambda - r_1 + v)^{-1}(e^{-vt} - e^{(\lambda - r_1)t}).$$

They examine three situations: where the system begins in equilibrium and then there is a single badger culling, where there is a sustained cull, but the testing frequency is changed, and where ther is no removal/restocking, but the badgers/cattle are still infecting eachother.

In order to provide numerical results, the equations are made dimensionless: $r_1 t = \tau$ and $l = \lambda/r_1$. The equations used for the population of cattle are made dimensionless by dividing through by $C^* = \frac{s_1 + a_{12}B^*}{r_1 - a_{11}}$.

```
In [5]: def plot_curves(curves, params_list, title):
        plt.figure(figsize=(12, 6), dpi=300)
        plt.hold()
        for curve, params in zip(curves, params_list):
            plt.plot(ts, curve, label=r'$k=\{k\}, I_B=\{I_B\}, l=\{l\}$'.format(*params))

        plt.xlabel(r'$\tau$')
        plt.ylabel(r'$C(\tau)/C^*$')
        plt.title(title)
        plt.legend()

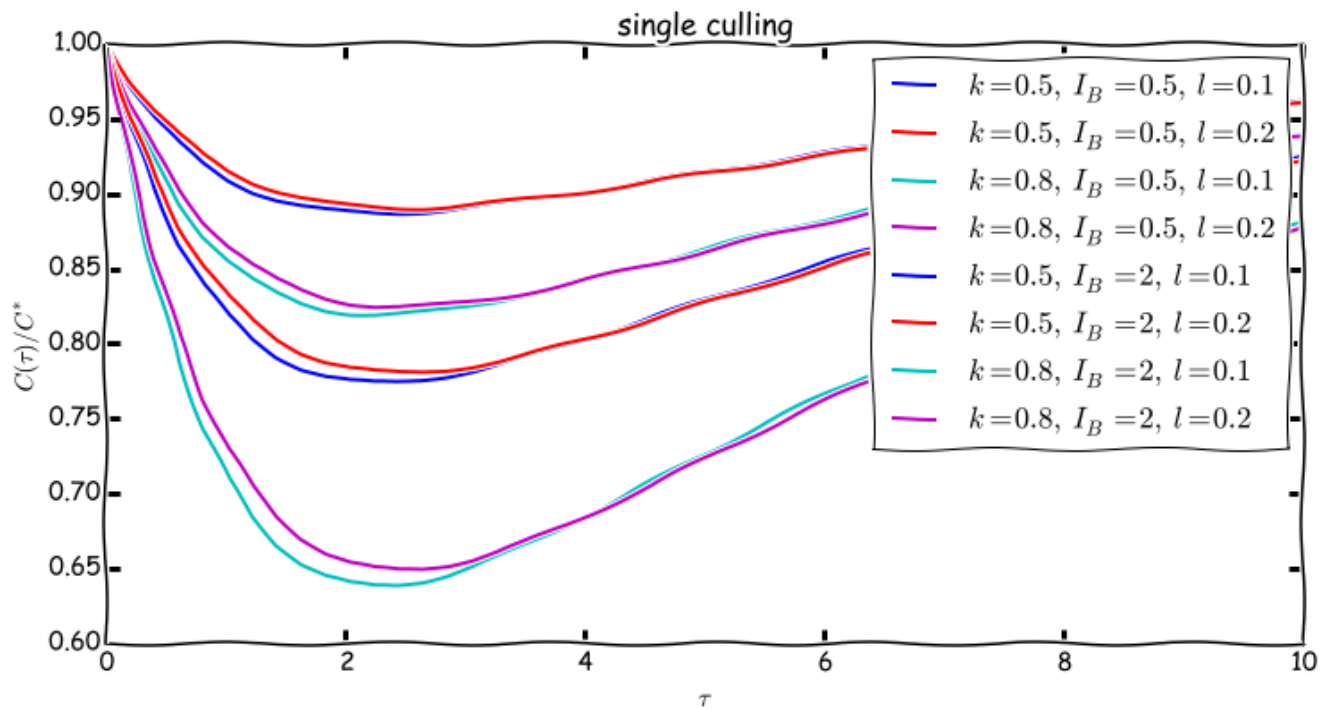
vr1 = 0.17
```

```
ts = linspace(0, 10)
params_list = [[.5, .5, .1],
               [.5, .5, .2],
               [.8, .5, .1],
               [.8, .5, .2],
               [.5, 2, .1],
               [.5, 2, .2],
               [.8, 2, .1],
               [.8, 2, .2]]
```

Dimensionless Graph of Effect of Single Culling

```
In [7]: def singleCull(t, k, i, l):
        num = k * i * (1. - l)
        denom = (1 + i)*(1. - l - vr1)
        return 1. - (num/denom) * (exp(-vr1 * t) - exp((1-l.) * t))

curves = [[singleCull(t, *params) for t in ts] for params in params_list]
plot_curves(curves, params_list, 'single culling')
```



In order for this system to have epidemic growth, the matrix $\vec{B} = \vec{R} - \vec{A}$ has to have one negative eigenvalue corresponding to exponential growth.

Proposal

After reading the paper and some other materials for clarification, I have two ideas that I would like to pursue (I'll just do whichever you select):

1. Data science on the other species and see if the model holds
2. Try doing calculations with a more complicated badger culling model

For number 1, there is extensive data made available by the British government. It can be found [here](https://www.gov.uk/government/publications/incidence-of-tuberculosis-tb-in-cattle-in-great-britain) (<https://www.gov.uk/government/publications/incidence-of-tuberculosis-tb-in-cattle-in-great-britain>).

Badger culling has been in the news lately and has been reported as ineffective and cruel. I'd like to look into more recent research (both by Cox and other research groups) and see how much the model is complicated by less trivial culling models (and to what extent the model predicts real-world ineffectiveness).

Sorry that this was so late. I had so much work last week, that I couldn't get started on this until friday night.