

Optimization and fitting

ed: Eli Gurarie

QERM 598 - Lecture 9

March 5, 2008

Meet Cassava Mosaic Disease (CMD) ¹

Cassava is a primary food source for more than 250 million Africans (about 40% of the continent's population). Its starchy root is a substantial portion of the diet of nearly 600 million people worldwide. In fact, cassava is relatively easy to grow in drought conditions and its root can stay in the ground for up to two years.



Infected Cassava leaf



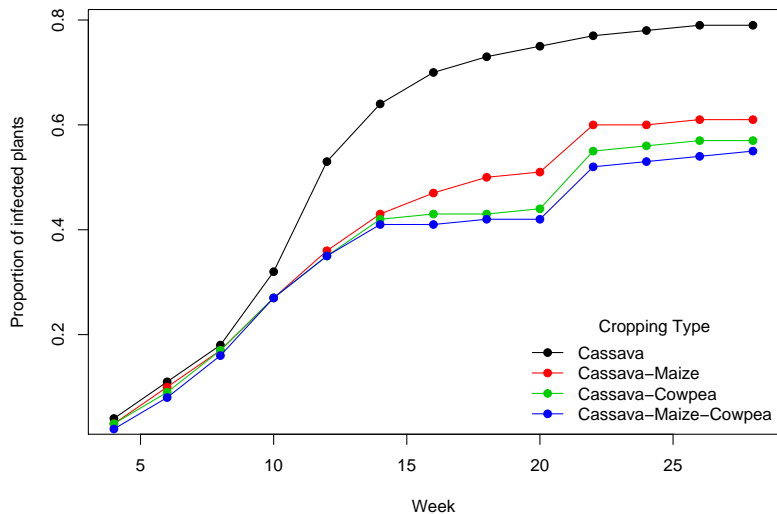
Bemisia tabaci adult.

Cassava mosaic disease (CMD) is the most important disease of cassava in Africa and the Indian subcontinent. The disease is caused by cassava mosaic viruses of the genus *Begomovirus* (family Geminiviridae) that are transmitted by the whitefly, *Bemisia tabaci*.

¹Text and data lifted wholesale from:

http://www.apsnet.org/Education/AdvancedPlantPath/Topics/RModules/Doc1/08_CMD.html

Mixed Cropping Type as a treatment



Sigmoidal Curves

Exponential Regression Model aka von Bertalanffy growth:

$$f(t|a, b, \lambda) = a - be^{-\lambda t} \quad (1)$$

Weibull Regression Model:

$$f(t|a, b, \lambda, h) = a - be^{-\lambda t^h} \quad (2)$$

Logistic Regression Model:

$$f(t|a, b, \lambda) = \frac{a}{1 + be^{-\lambda t}} \quad (3)$$

Richards Regression Model:

$$f(t|a, b, \lambda, h) = \frac{a}{(1 + be^{-\lambda t})^h} \quad (4)$$

Preparing to use `optim()`

Define function:

```
Model1 <- function(t,a,b,lambda)
  return(a-b*exp(-lambda*t))
```

Define measure to minimize (e.g. sum of squares):

```
SS <- function(p)
{
  a<-p[1]
  b<-p[2]
  lambda<-p[3]
  return(sum((mydata-mymodel(Week,a,b,lambda))^2))
}
```

Important! Note that this function **MUST** take a **VECTOR** of parameters and return a **SINGLE** value!

Preparing to use `optim()`

Define function:

```
Model1 <- function(t,a,b,lambda)
  return(a-b*exp(-lambda*t))
```

Define measure to minimize (e.g. sum of squares):

```
SS <- function(p)
{
  a<-p[1]
  b<-p[2]
  lambda<-p[3]
  return(sum((mydata-mymodel(Week,a,b,lambda))^2))
}
```

Important! Note that this function **MUST** take a **VECTOR** of parameters and return a **SINGLE** value!

Using `optim()` to estimate parameters

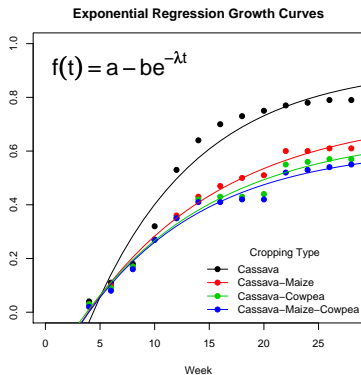
Using `optim`:

```
p0 <- c(1,1,.1) # initialized parameters
optim(p0, SS)    # optimize!
```

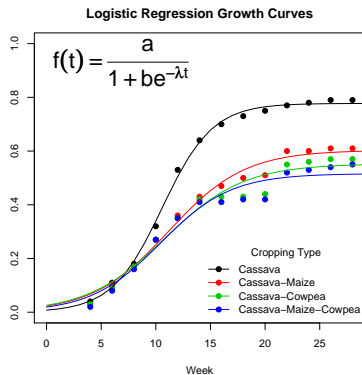
Output:

```
$par
[1] 0.1535428 0.5099559 0.4193734
$value
[1] 0.2040492
```

Sigmoidal Curves

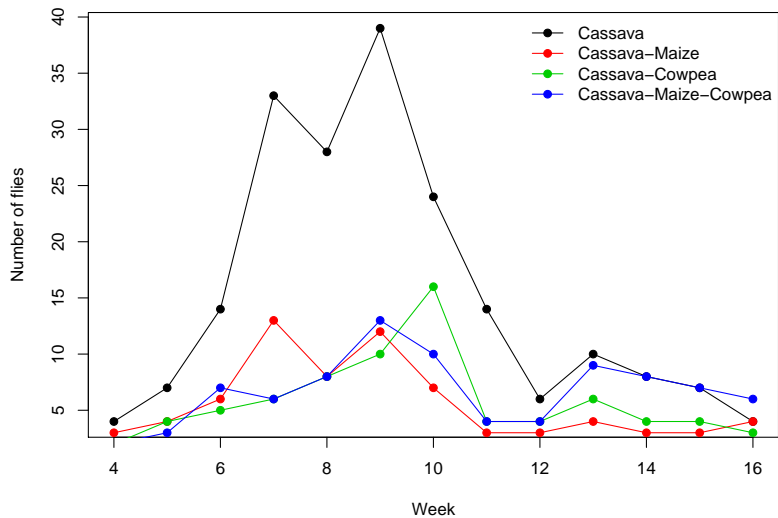


	C	C.M	C.P	C.M.P
a	0.92	0.74	0.67	0.62
b	1.43	1.01	0.92	0.89
λ	0.10	0.08	0.08	0.09

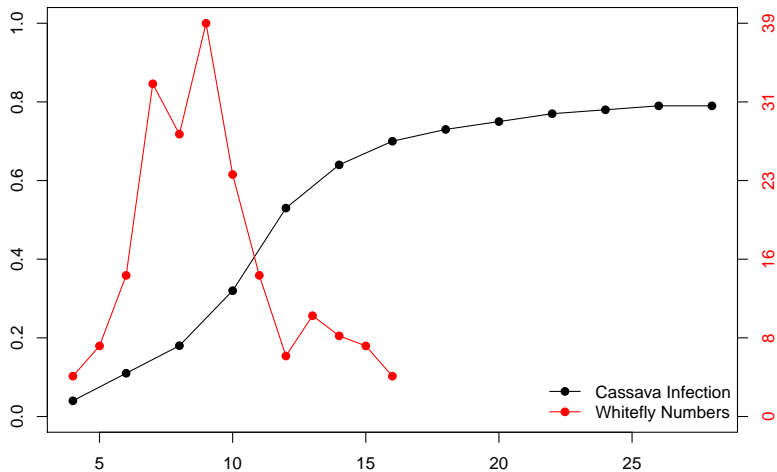


	C	C.M	C.P	C.M.P
a	0.78	0.60	0.55	0.52
b	100.72	25.15	20.60	26.35
λ	0.44	0.29	0.28	0.31

White fly numbers



White fly numbers



Simple Whitefly-Cassava Infection Model

$$\frac{dI}{dt} = a(1 - I)N \quad (5)$$

$$\frac{dN}{dt} = b(1 - I) - cN \quad (6)$$

where:

- I is proportion of infected plants
- N is number of whiteflies
- a is rate of infection per whitefly per susceptible plant
- b is rate of arrival of whiteflies per susceptible plant
- c is mortality rate of whiteflies

Convert to difference Equation

$$I_{t+1} = a(1 - I_t)N_t + I_t \quad (7)$$

$$N_{t+1} = b(1 - I_t) + (1 - c)N_t \quad (8)$$

in R-code:

```
I <- rep(0,tmax)
N <- rep(0,tmax)
for(i in 2:tmax)
{
  I[i] <- (1-I[i-1])*N[i-1]*a + I[i-1]
  N[i] <- (1-I[i-1])*b + (1-c)*N[i-1]
}
```

Convert to difference Equation

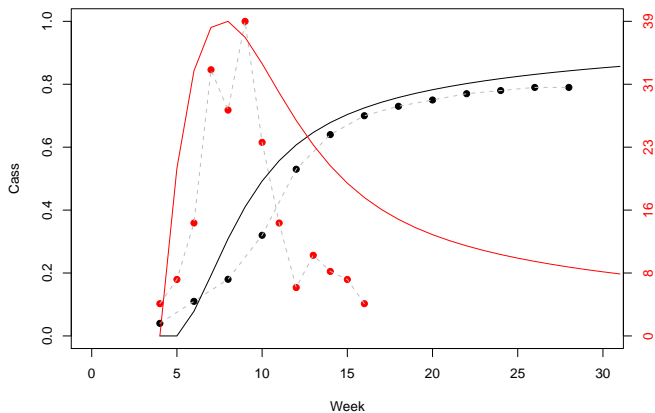
$$I_{t+1} = a(1 - I_t)N_t + I_t \quad (7)$$

$$N_{t+1} = b(1 - I_t) + (1 - c)N_t \quad (8)$$

in R-code:

```
I <- rep(0,tmax)
N <- rep(0,tmax)
for(i in 2:tmax)
{
  I[i] <- (1-I[i-1])*N[i-1]*a + I[i-1]
  N[i] <- (1-I[i-1])*b + (1-c)*N[i-1]
}
```

Fitted Model (using optim)



$$a = 0.15; b = 0.51; c = 0.42$$

In-class

- Obtain estimates for the other cropping combinations
- Is the model good?
- How could it be refined?
- Do we learn anything from the lack of fit?