Population Modeling of Smartphone Sales

Hugh Higinbotham, Qiao Liu, Kandi Lou

June 6, 2015

1 Abstract

In this paper, we attempt to adapt the competitive species form of the Lotka-Volterra equations to model the dynamics between Apple and Android smartphone sales. We start with examining the simple Logistic Growth model and proceed to adapt the Lotka-Volterra equations in several different ways in order to determine how the two companies affect one another. We conclude by discussing which model seems to work the best and how we would improve upon this model in the future.

2 Introduction

By the first quarter of 2015, smartphones carrying either an IOS or Android operating system has constituted 96.3 percent of global market share. Because their huge percentage of the market share and dramatic growth of both systems, the competition between these two major operating systems will determine the future trends for global smartphone market. In our paper, we will answer the following questions: What is the future trends for smartphones carrying these two systems? When will the market reach its maximum capacity? And what is the saturation point for both systems? These are all real world problems that is going to affect the future for global smartphone market. Since Apple and Android are the two dominant forces in the market and they are competing over the resource of consumers, our idea is to use population models to predict the general trends of smartphone sales. Specifically, since we are considering the competition between two market entities, we are attempting to adapt the Lotka-Volterra equations for competitive species:

$$\frac{dN}{dt} = r_N N(1 - \frac{N+\alpha_P P}{K_N})$$

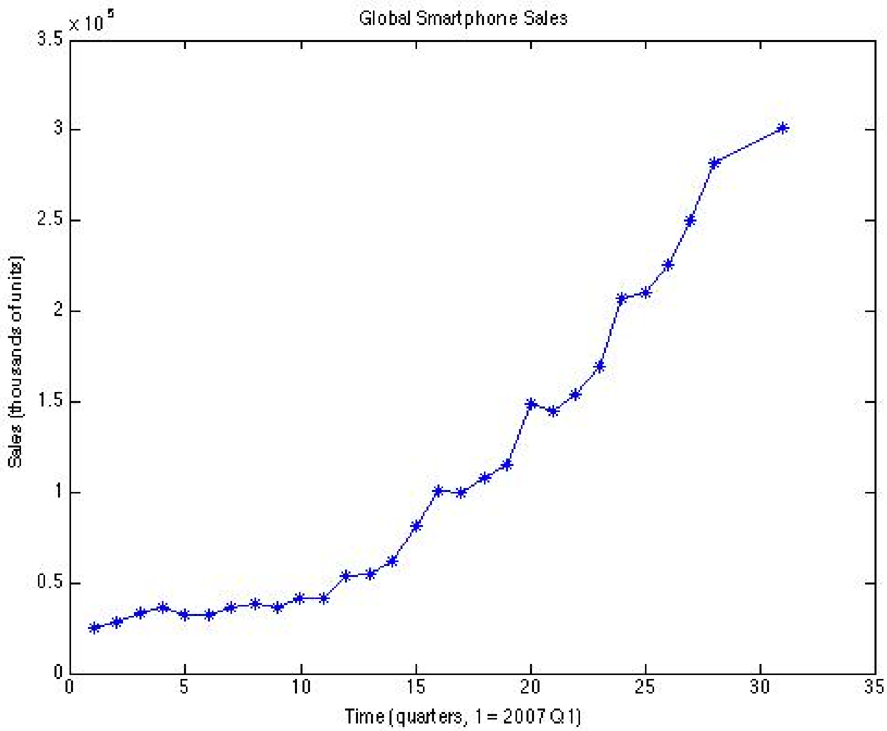
$$\frac{dP}{dt} = r_P P(1 - \frac{P+\alpha_N N}{K_P})$$

Where N and P are the 2 populations, r_N and r_P are the rates of growth, K_N and K_P are the carrying capacities of the environment, and α_P and α_N are the coefficients representing how one species' population affects the other. This is a modification of the Logistic Growth population model, given by

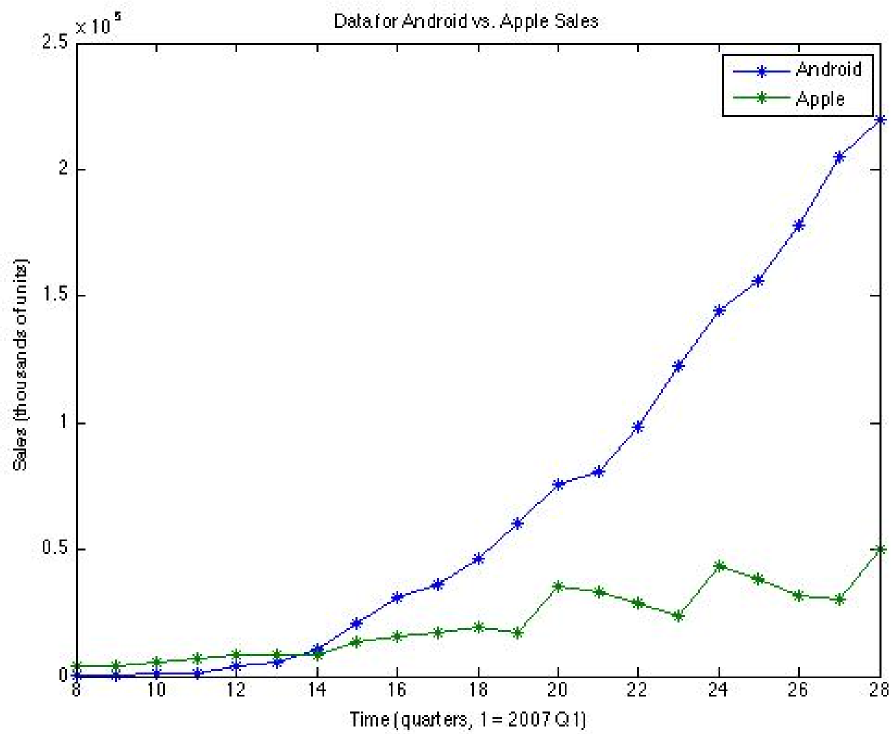
$$\frac{dQ}{dt} = rQ(1 - \frac{Q}{K})$$

for a population Q, where r is the rate of population growth and K is the carrying capacity of the population's environment.

We got our data on the phone sales from Gartner, an extensive database for marketing research (see references), as follows:



Global Smartphone Sales
x 10^5
Sales (thousands of units)
Time (quarters, 1 = 2007 Q1)
0
0.5
1
1.5
2
2.5
3
3.5
0
5
10
15
20
25
30
35



Note that the time axis is given in quarters, where 1 corresponds to 2007 quarter 1 (when our data begins). The Y axis is sales, given in thousands of units sold.

3 Initial Assumptions

In order to use this model, we had to make a number of broad assumptions about the smartphone market. Since Apple and Android dominate the market to such a degree, we are only considering these two types of smartphones and ignore the effects of all others in order to reduce complexity. Additionally, by only considering the variables of the rates of increase, the carrying capacities, and the effects of one company on the other, we ignore other potentially confounding factors and events (such as the emergence of new technologies, a radical change in the global economy, etc.) that would complicate the situation beyond the scope of our chosen model.

4 Model

4.1 Logistic Growth Model

We started with a simple logistic growth model, since the Lotka-Volterra is a modification of this model, and used it as a baseline to compare how well the later models fit the data.

For a population Q, solving the Logistic Growth Equation yields

$$Q(t) = \frac{KQ_0e^{rt}}{K+Q_0(e^{rt}-1)}$$

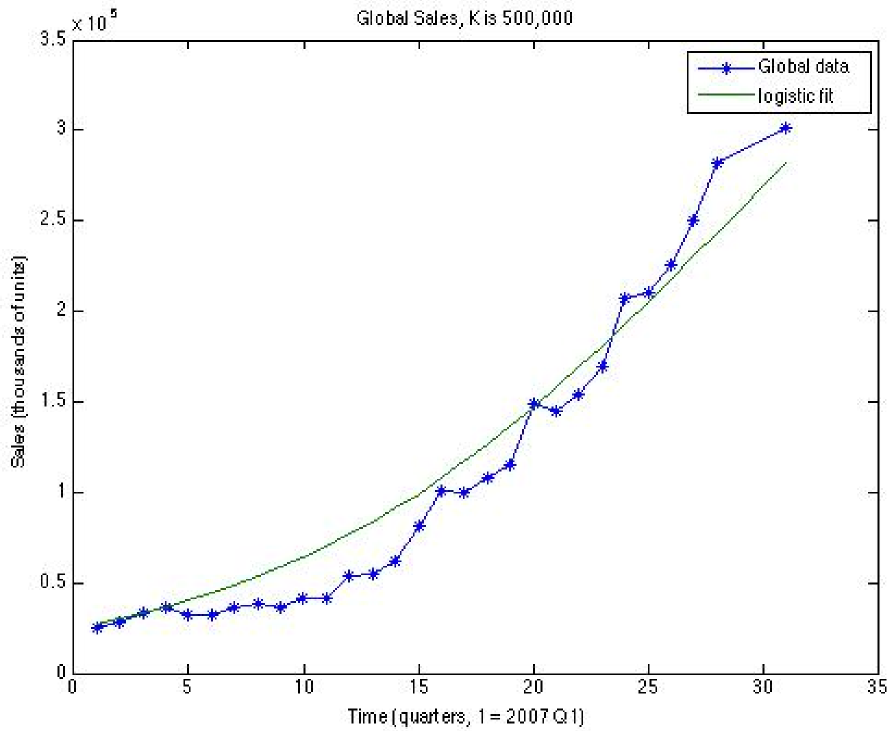
Where Q_0 is the population of Q at t = 0. For simplicity, we used this integrated form in calculating logistic fits to our data.

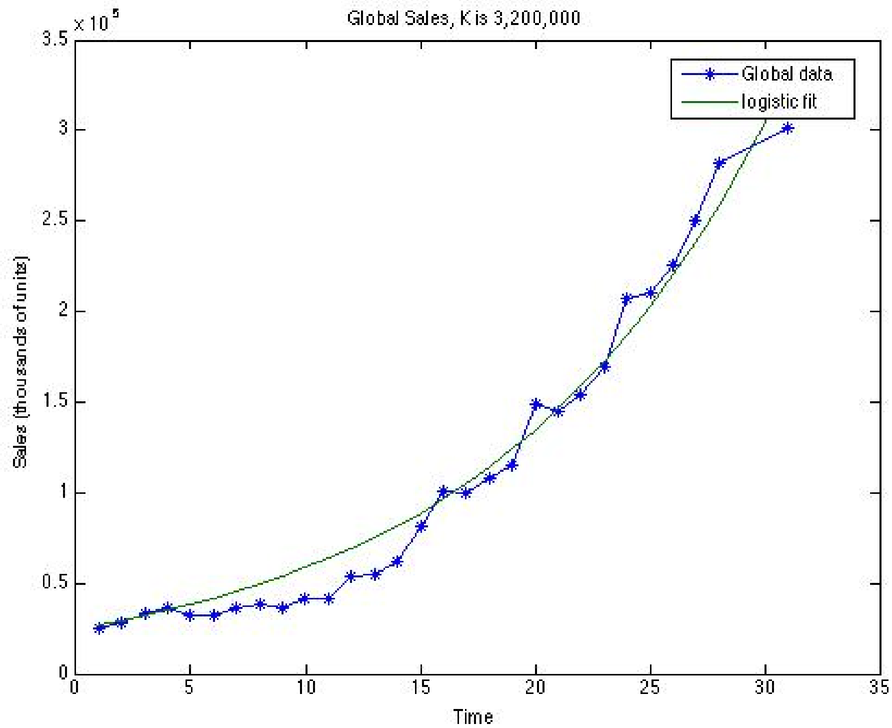
4.1.1 Additional Assumptions

This model assumes that phone sales can be modeled as the growth of a population whose environment has a specific carrying capacity K. We assume the rate of increase is constant in time and that there is no effect from one ?species? (or phone company sales) on the other.

4.1.2 Data Fitting

We first used the data from global smartphone sales to determine an estimate for the carrying capacity K and the rate of increase of smartphone sales as a whole. Since the number of sales of all mobile phones globally has been nearly constant at around 500,000 thousand units per quarter (data also received from Gartner), we assumed this as the carrying capacity of the system. I.E. the number of global smartphone sales will cap out at the number of total cell phone sales, which would occur when smartphones completely replace all other types of cell phones. To evaluate the validity of this assumption, we compared the error in a logistic fit with this assumed K to a logistic fit that optimizes for K. This error we calculate as the sum of the squares of the differences of our model's prediction and the data from Gartner at every point.


Global Sales, K is 500,000
x 10^5
Global data
logistic fit
Sales (thousands of units)
Time (quarters, 1 = 2007 Q1)
0
0.5
1
1.5
2
2.5
3
3.5
0
5
10
15
20
25
30
35



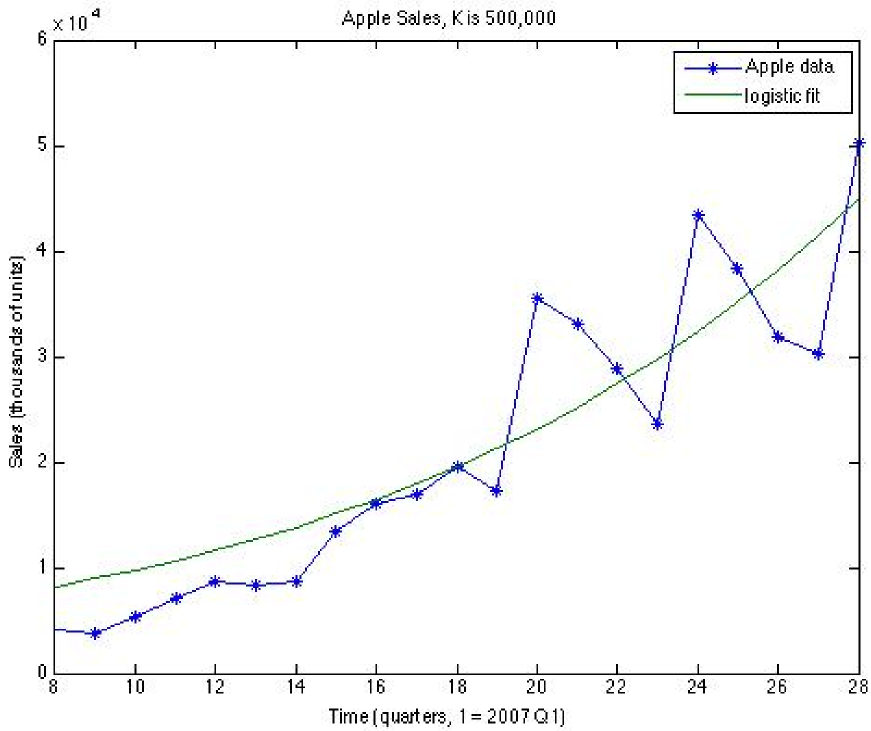
The error of the optimized K is $4.8 * 10^9$ and the assumed K is $9.28 * 10^9$, which is measured by the sum over all quarters of the squared difference of the logistic model vs. the real-world data. While optimizing for K halves the total error in the model, in order to do so it assumes that K is approximately 3.2 billion units sold per quarter. Additionally, initial tests on Android and Apple logistic fits showed that they are better fit by assuming a lower K than this. As such, we considered this optimized value of K an unrealistic artifact and stuck with our assumption of K = 500,000 thousands of units given that it better suits the reality of our modeling situation. For this K, the optimized rate we found is r = 0.1034 thousand units per quarter.

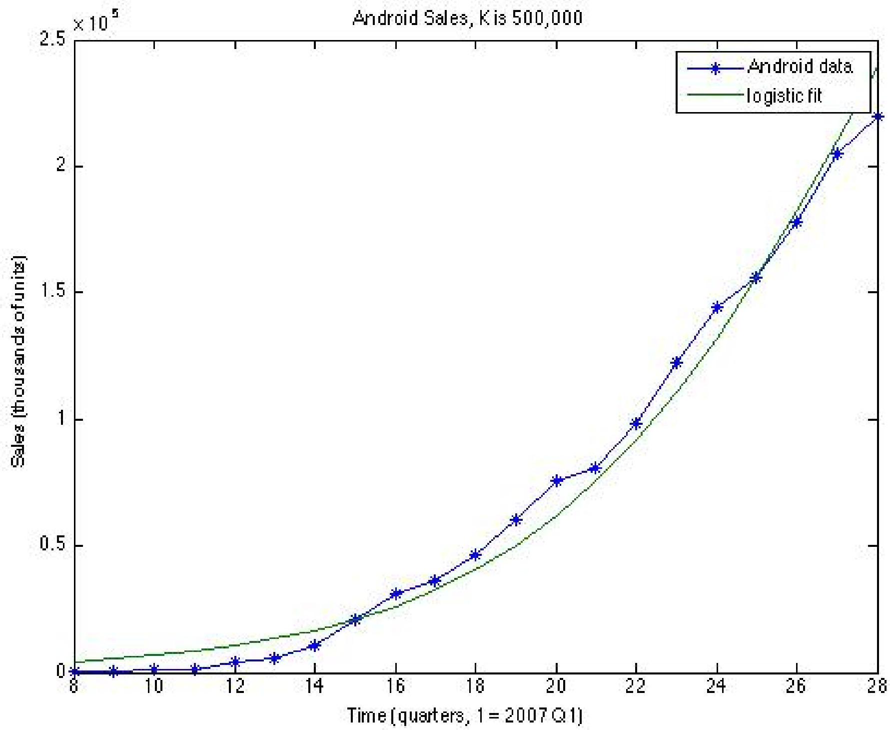
Next, we seek the optimal the growth rate for both Android (N) and Apple (P) with logistic growth fits:

$$N(t) = \frac{KN_0e^{r_Nt}}{K+N_0(e^{r_Nt}-1)}$$

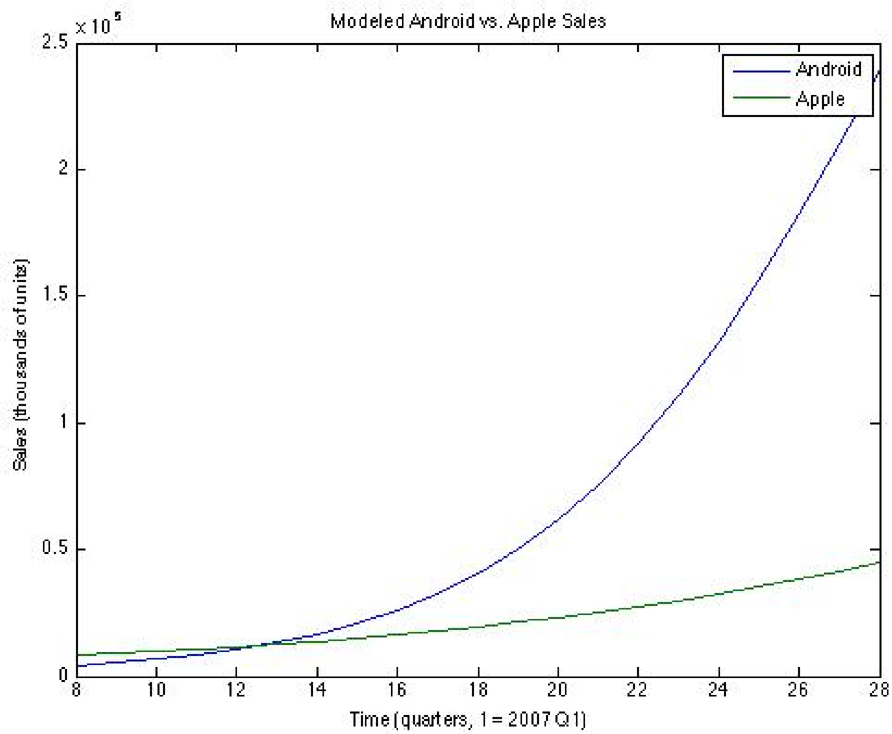
$$P(t) = \frac{KP_0e^{r_Pt}}{K+P_0(e^{r_Pt}-1)}$$

Where r_N and r_P are the rates of growth for each company's sales, and K is assumed to be the same as in our global approximation, as we are treating each company in isolation. To optimize, we minimize the same form of error function as we did in the global case to get the following fits:





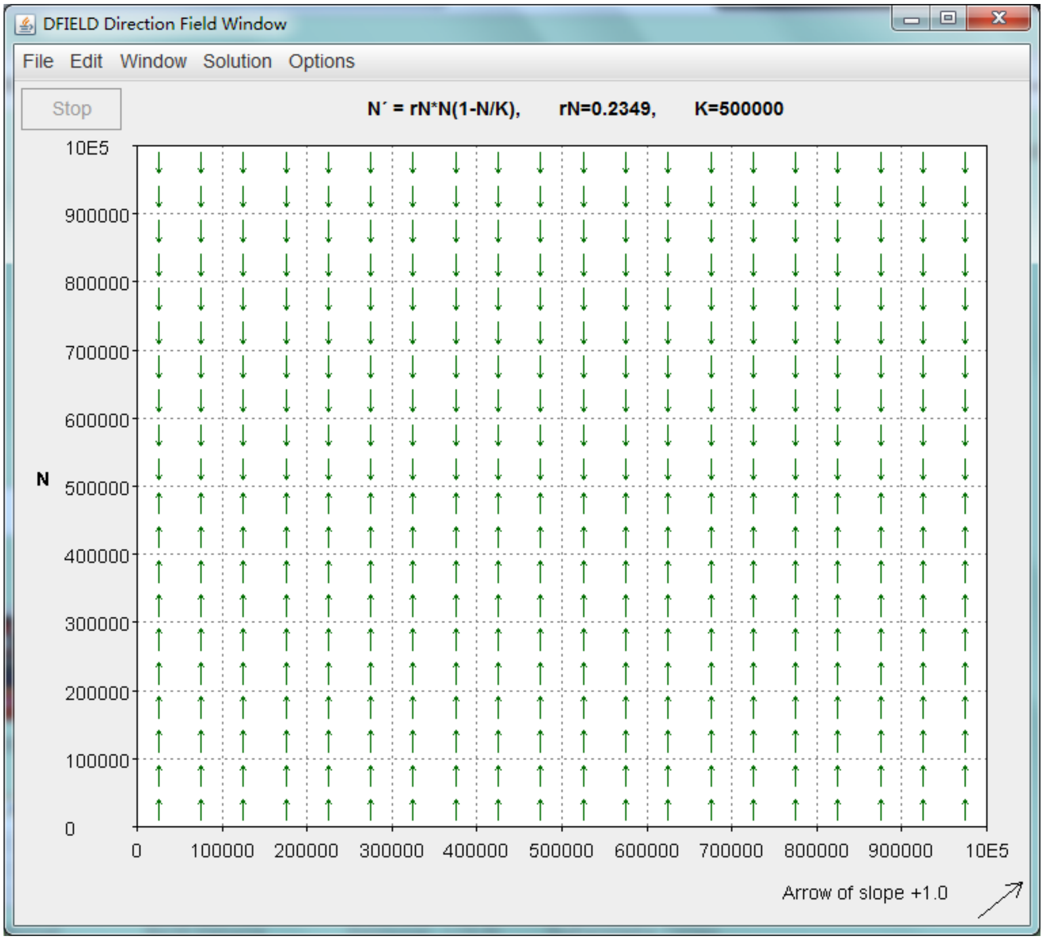
We found that the optimal growth rate for Android is $r_N = 0.2349$, and that for Apple it is $r_P = 0.0888$. Here are the two models side by side:



Errors:
Apple error: 7.2969e+08
Android error: 1.4764e+09
Combined Error: 2.2061e+09

4.1.3 Stability Analysis

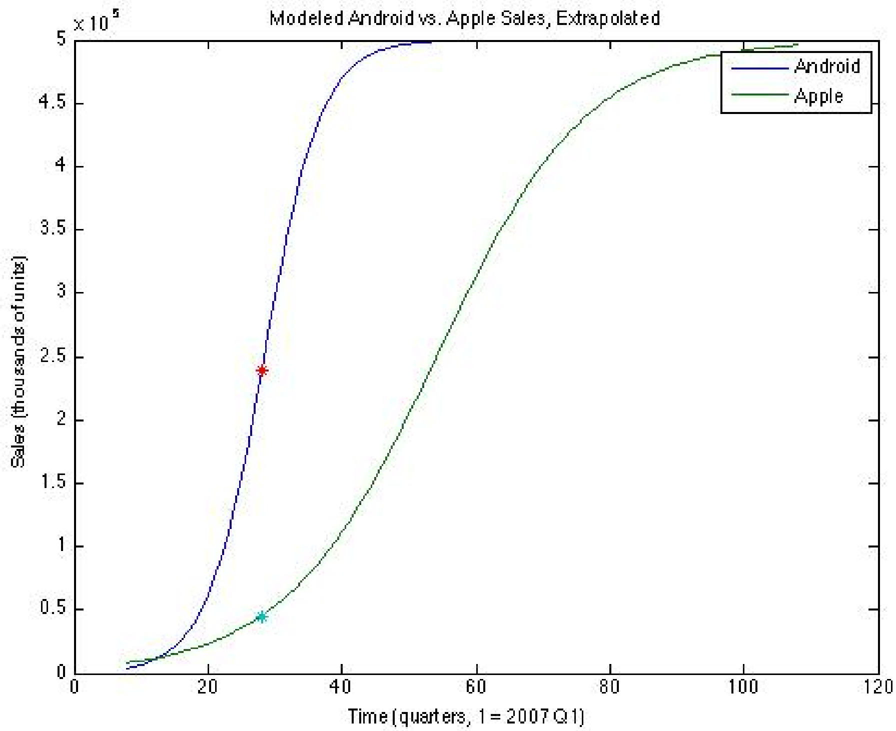
For our stability analysis in this paper, we used an open source phase plane generator put online by Rice University's departments of mathematics (see references for link).



From the direction field graph we can see that Android sales will grow to K thousand units and reach equilibrium. Since this is true for any value of r_N, Apple sales would exhibit similar behavior, as except for the rate it is identical to the Android equation. This fits with the expected behavior of a logistic growth model with carrying capacity K.

4.1.4 Extrapolating the Data

To extrapolate this model, we extended the time period to include 100 total time values (quarters) which corresponds to 25 years. Since we have a nice integrated form for the Logistic equation, we were able to simply plug in the extra data in this case in order to observe the asymptotic behavior.



The marked points on the graph indicate when our data ends and the prediction begins.

4.1.5 Efficacy

While a visual inspection suggests that these fits match the general trend of the data decently well, examining the extrapolation and stability of this model shows that it doesn't fit with our scenario since we considered both companies as completely independent, this model predicts that they ail soon both asymptotically approach the carrying capacity K of the market. As such, our next goal is to find a way to couple the equations such that its behavior at long times is more reasonable given the scenario.

4.2 2 Species Lotka-Volterra with Predation

Examining the logistic fits for the data, we see that the rates of increase for Apple and Android phone sales are both much larger than the rate we found for the global market. If we keep the assumption that the rates of phone sales would equal that of the global market independent of effects between the different populations, we cannot use the competitive Lotka-Volterra equations due to the fact that it assumes only inhibitory effects from one population on another (I.E. α_P and $\alpha_N \geq 0$). This would result in the optimal value of alpha being 0, and we would recover the logistic fit to the global market from this equation. As such, we tried fitting for a Lotka-Volterra model that allows alpha coefficients to be negative, corresponding to mutualism (where one species growing increases that of another species). This corresponds to greater Android sales increasing Apple sales and vice-versa.

4.2.1 Additional Assumptions

The key assumption of this model is that $r_N = r_P = r$ where r is the rate at which the global market increases. This supposes that the only effects on the growth rate of different 'species' are due to the other population (I.E. that Android and Apple sales only grow at different rates because of how they affect one another). While this is a very naive model, we hoped to glean some general information about how the companies affect one another when excluding all other factors.

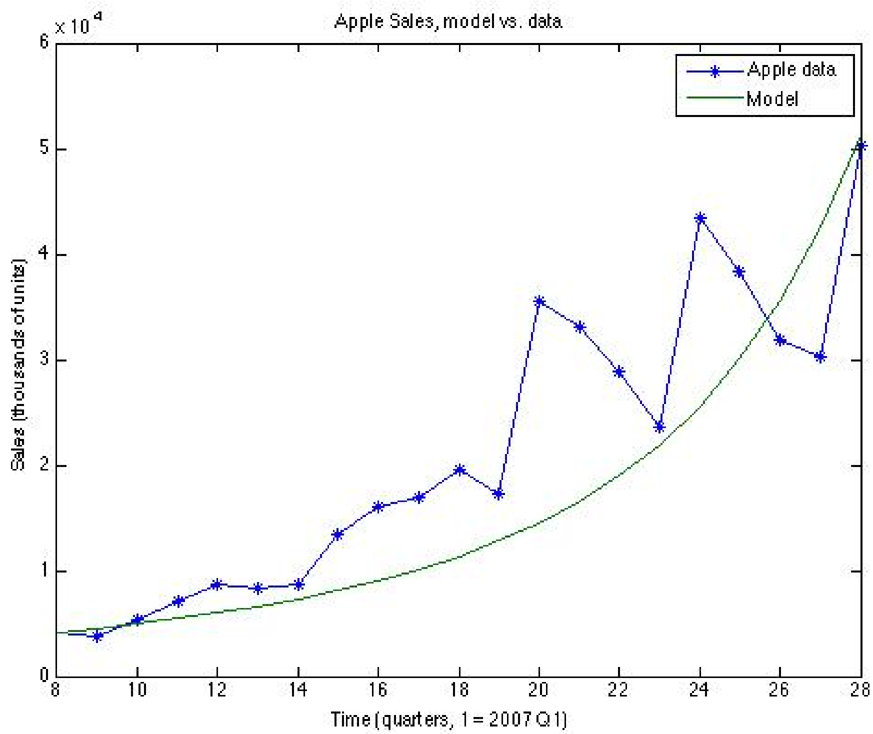
4.2.2 Data Fitting

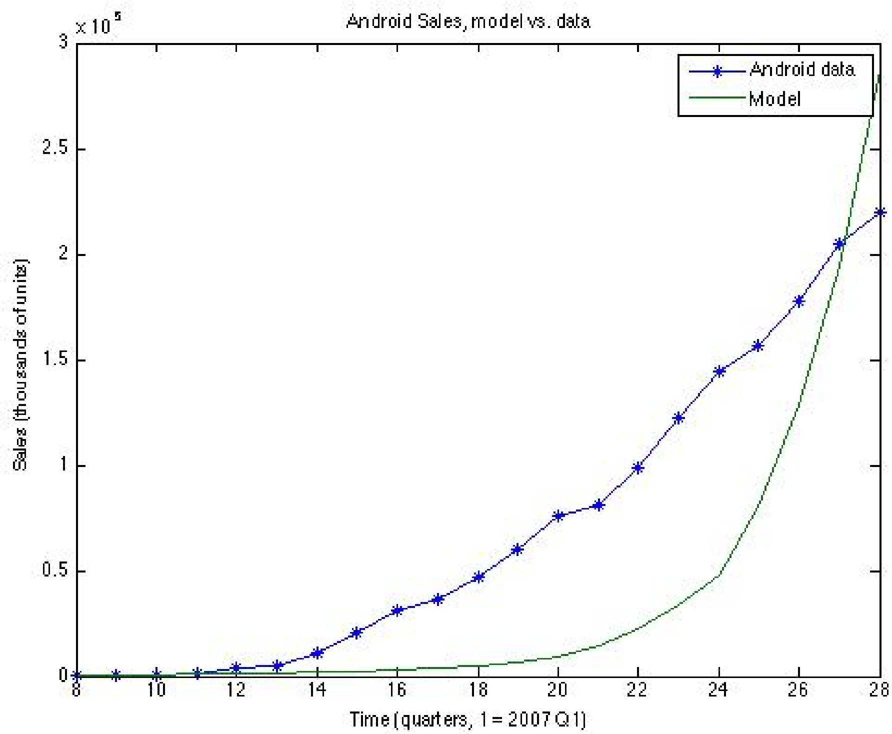
We found these fits by iterating through the discrete Lotka-Volterra equations and minimizing the sum of the squares of the differences between our model?s values and the data.

$$N(t + \Delta t) = N(1 + r\Delta t N(1 - \tfrac{N+\alpha_P P}{K})$$

$$P(t + \Delta t) = P(1 + r\Delta t P(1 - \tfrac{\alpha_N N+P}{K})$$

Where Δt is one quarter for our data. Minimizing the error with respect to the alpha coefficients yielded alpP = -2.648 and alpN = -65.9201.


Apple Sales, model vs. data
x 10^4
Apple data
Model
Sales (thousands of units)
Time (quarters, 1 = 2007 Q1)



As is easily visible from the plots, this model does significantly worse than a pure logistic fit.

Errors:

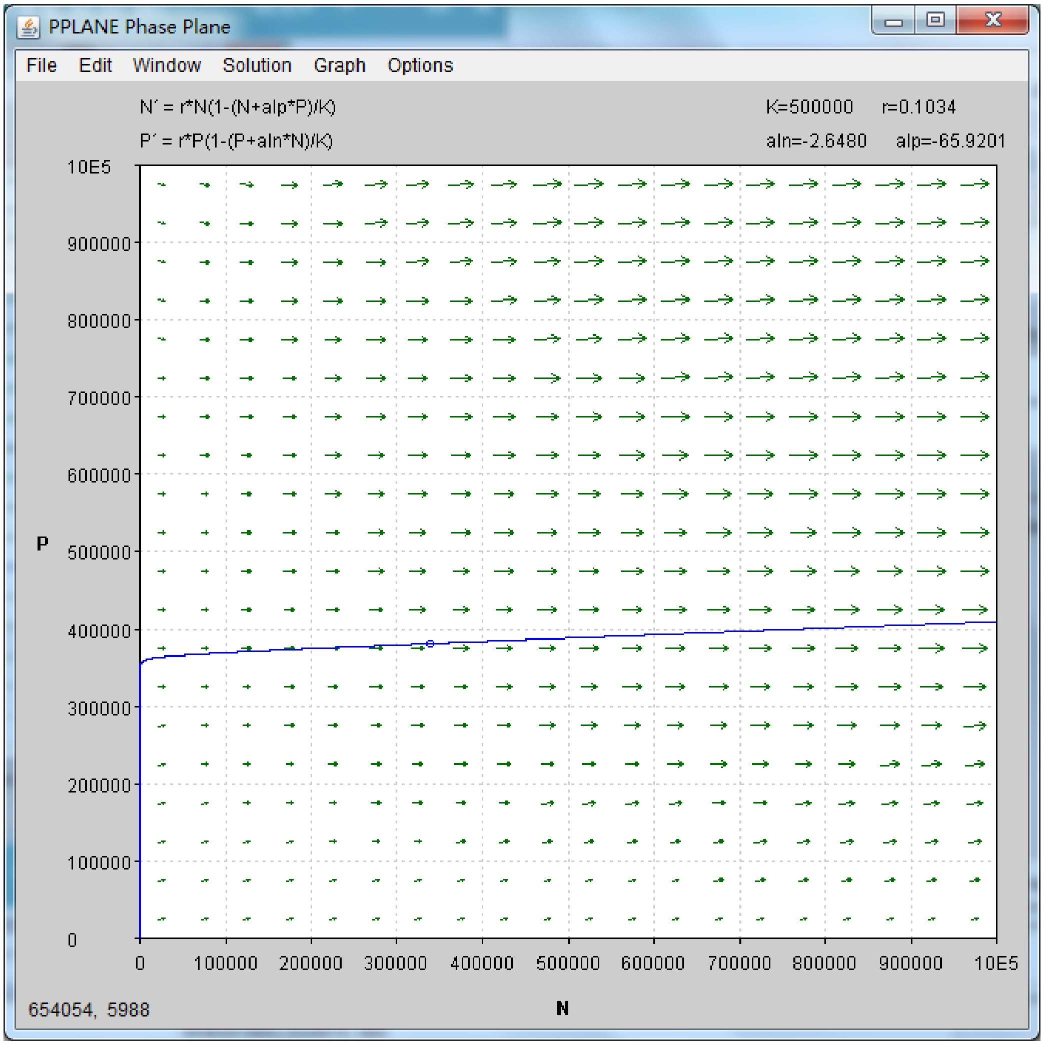
Apple error: 1.5863e+09

Android error: 5.1929e+10

Combined error: 5.3515e+10

4.2.3 Stability Analysis

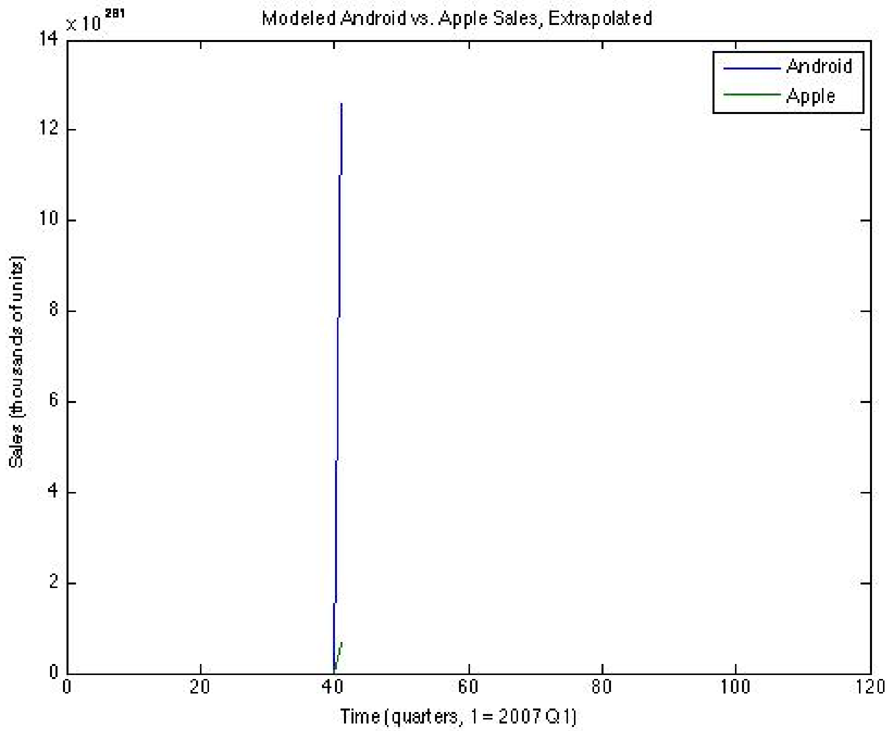
Below is the phase plane for the Lotka-Volterra model with predation, with $r_N = r_P = r = 0.1034$ where the coefficients $\alpha_P = -65.92$ and $\alpha_N = -2.648$ were the values we found in the above data fitting:



Note that there are no stable solutions for this model assuming negative alpha coefficients, as the carrying capacity is no longer an equilibrium point.

4.2.4 Extrapolating the Data

Attempting to extrapolate from the data by iterating this model yielded the following graph:



Wherein the populations explode to infinity almost immediately after the region in which we have data.

4.2.5 Efficacy

This model was not only poorly able to fit the data compared to a simple logistic fit, but its asymptotic behavior shows that it is an entirely unrealistic model to adopt for the data. This implies that independent rates of increase for Apple and Android sales are key factors in addition to how much the two companies sales affect one another.

4.3 Competitive Lotka-Volterra with independent rates

In light of the failure of the previous model, we concluded that a competitive Lotka-Volterra model of the situation in which the different species had different rates of increase would be better suited to the data. As such, we fitted these equations independently with their respective rates and alpha coefficients as free parameters, using the same form of error function as in the above models:

$$\frac{dN}{dt} = r_N N(1 - \frac{N+\alpha_P P}{K})$$

$$\frac{dP}{dt} = r_P P(1 - \frac{P+\alpha_N N}{K})$$

4.3.1 Additional Assumptions

While we still assume that the rates and alpha coefficients are constant in time, we have abandoned the assumption that the rates of each species are related to the total rate of smartphone sales increase.

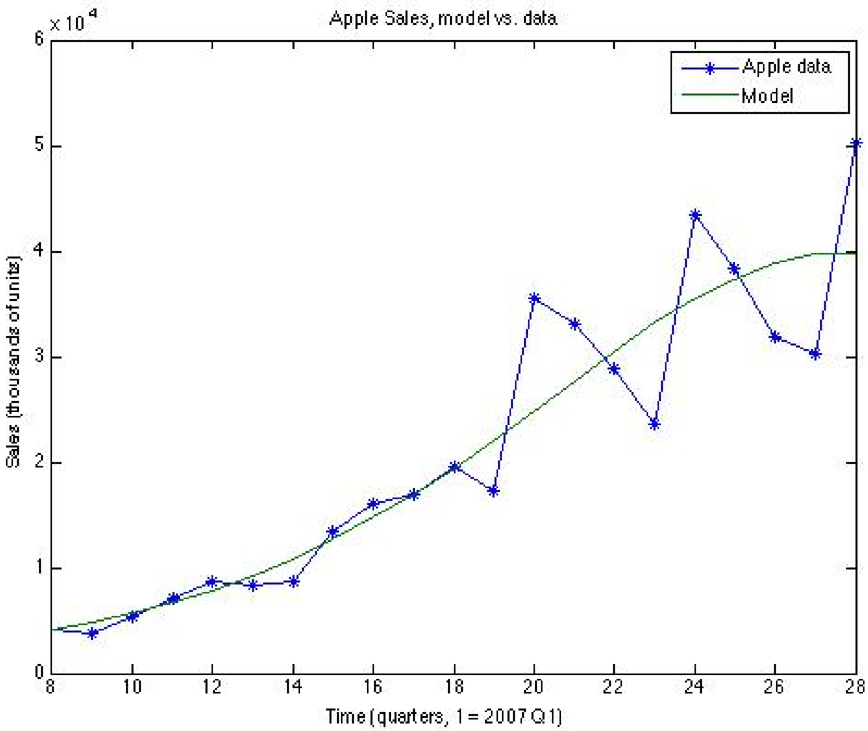
4.3.2 Data Fitting

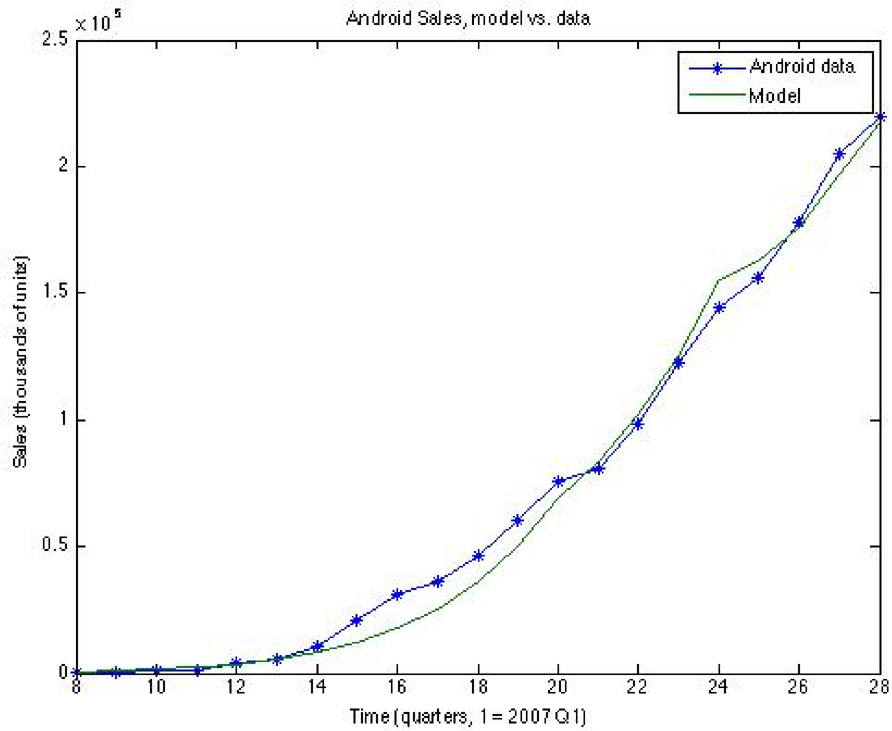
We optimized the data to reach minimum error, and in doing so got the growth rate, using the discrete form of the equations as before:

$$N(t+\Delta t) = N(1 + r_N \Delta t N(1 - \frac{N+\alpha_P P}{K})$$

$$P(t+\Delta t) = P(1 + r_P \Delta t P(1 - \frac{P+\alpha_N N}{K})$$

We found that the rate for Android $r_N = 0.5824$, and the growth rate for Apple $r_P = 0.1816$, $\alpha_P = 6.961$, $\alpha_N = 2.246$. Below is the model and the data for both types of cell phones:

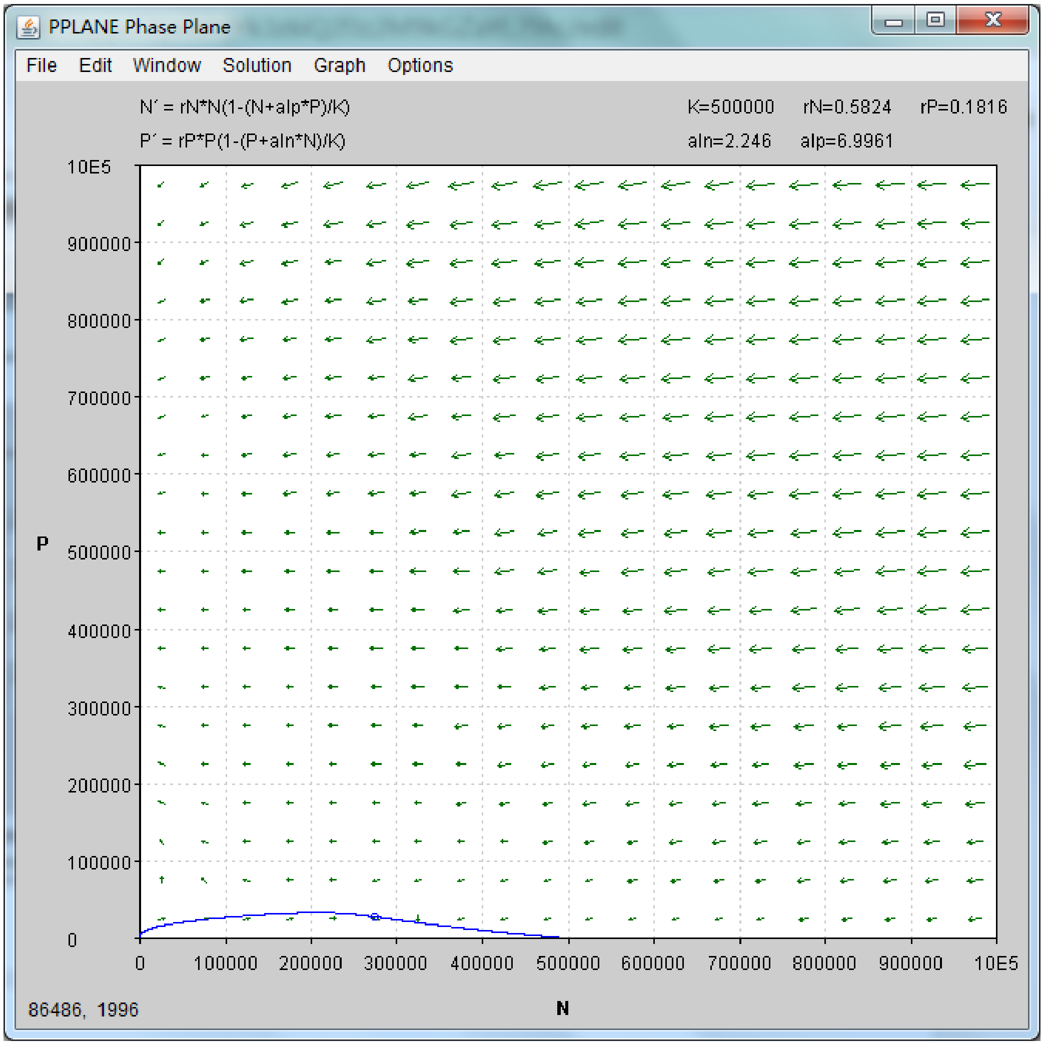

Apple Sales, model vs. data
x 10^4
Apple data
Model
Sales (thousands of units)
Time (quarters, 1 = 2007 Q1)
0
1
2
3
4
5
6
8
10
12
14
16
18
20
22
24
26
28



Errors:
Apple error: 5.8270e+08
Android error: 8.9723e+08
Combined error: 1.4799e+09

4.3.3 Stability Analysis

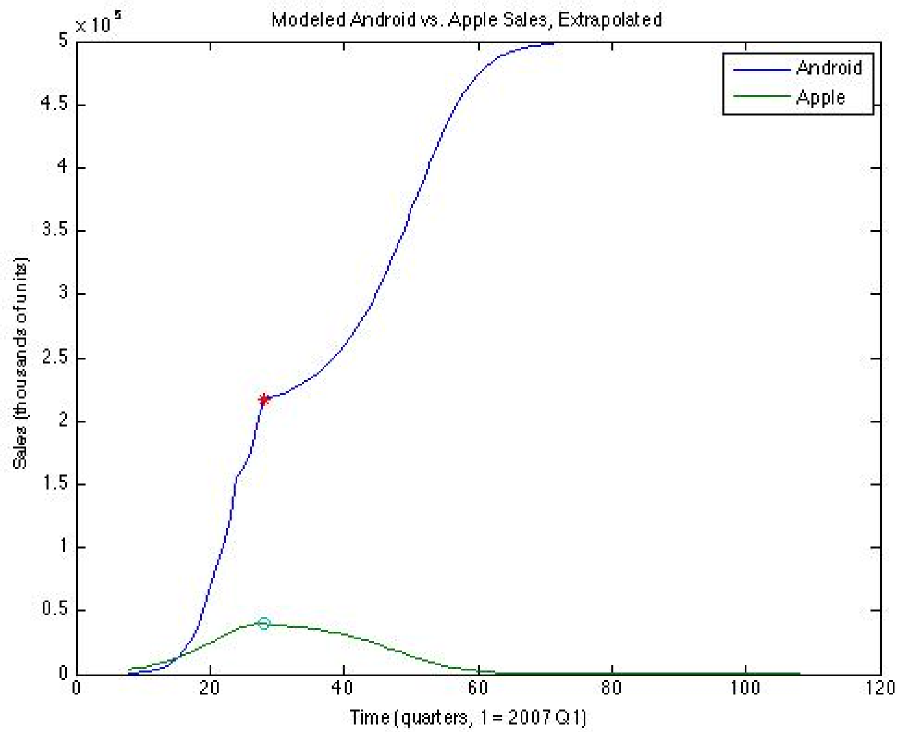
Below is the phase plane graph for the competitive Lotka-Volterra Model with independent rates for Android and Apple:



From the graph we can see that the extrapolation depends on the initial value: If we start at a point with more sales in Apple than Android, we would end up at the left of the graph, where Apple sales reach carrying capacity and Android sales decay to zero. But since the data we have start at a point where Android has significantly more sales than Apple, we will end up having Android grow to carrying capacity K and Apple decay to 0.

4.3.4 Extrapolating the Data

The data was extrapolated by iterating the discrete form of the equation for a total of 100 time steps. The marked points on each curve indicate where our data ends and the extrapolation begins.



4.3.5 Efficacy

While the overall error for this model is the best of any we have so far examined, looking at the stability analysis and extrapolation for this model suggests that it is not the best way to look at the smartphone market. The fact that the only non-zero equilibrium solutions occur when one company goes out of business while the other takes over the entire market is unrealistic, as there is no reason to think a monopoly in the smartphone market is inevitable or even likely. Additionally, extrapolating the Android and Apple data introduces very discontinuous behavior, especially for Android, at the point where we have no more data (as can be seen in the above extrapolation plot). This indicates that the competitive Lotka-Volterra equations with independent rates for Android and Apple, though it fits the data decently well, is probably not a good way to model smartphone sales.

4.4 Modified Lotka-Volterra

We constructed this new model such that it would exhibit more realistic asymptotic behavior. We expect both phones will reach a maximum capacity where their sum would end up to be our carrying capacity. While fitting for both the independent rates and [alpha] coefficients yielded better fits, this was in large part due to having more free parameters to vary. As such, we modified the equations to eliminate this extra parameter and predict more consistent and reasonable asymptotic behavior:

$$\frac{dN}{dt} = r_N N(1 - \frac{N+P}{K})$$

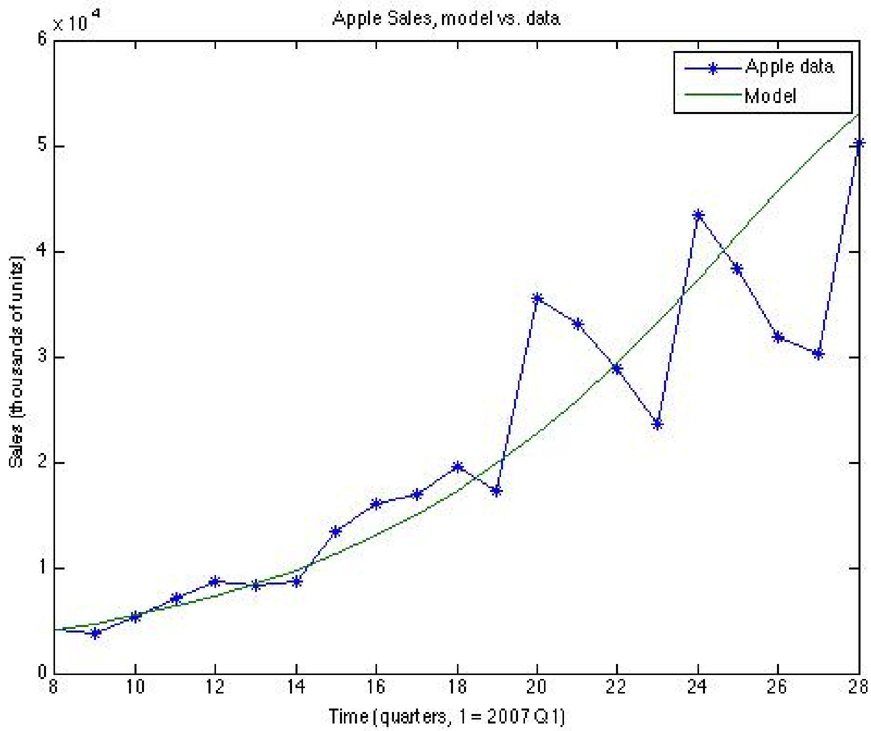
$$\frac{dP}{dt} = r_P P(1 - \frac{P+N}{K})$$

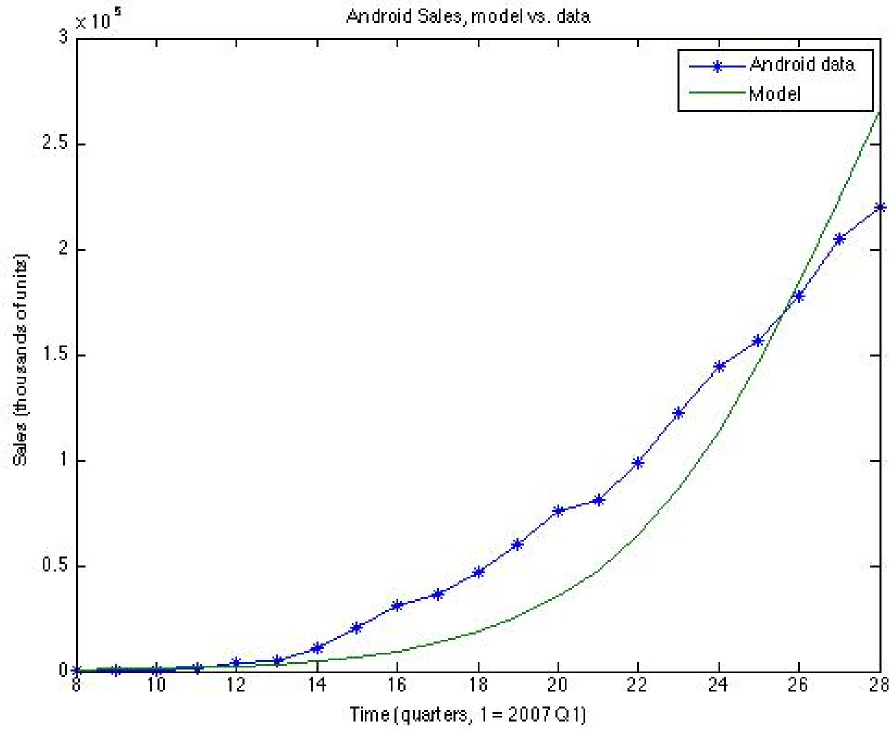
4.4.1 Additional Assumptions

This model assumes that Android sales and Apple sales affect each other equally (alphaP = alphaN = 1) and thus allows us to study the data solely in the context of the independent rates.

4.4.2 Data Fitting

By trying to get the minimum error, we get the optimized data to be: growth rate for Android rN = 0.4112, growth rate for Apple rP = 0.1596, the following graphs are the model we made with the actual data:


Apple Sales, model vs. data
Apple data
Model
Sales (thousands of units)
Time (quarters, 1 = 2007 Q1)
x 10^4
0
1
2
3
4
5
6
8
10
12
14
16
18
20
22
24
26
28

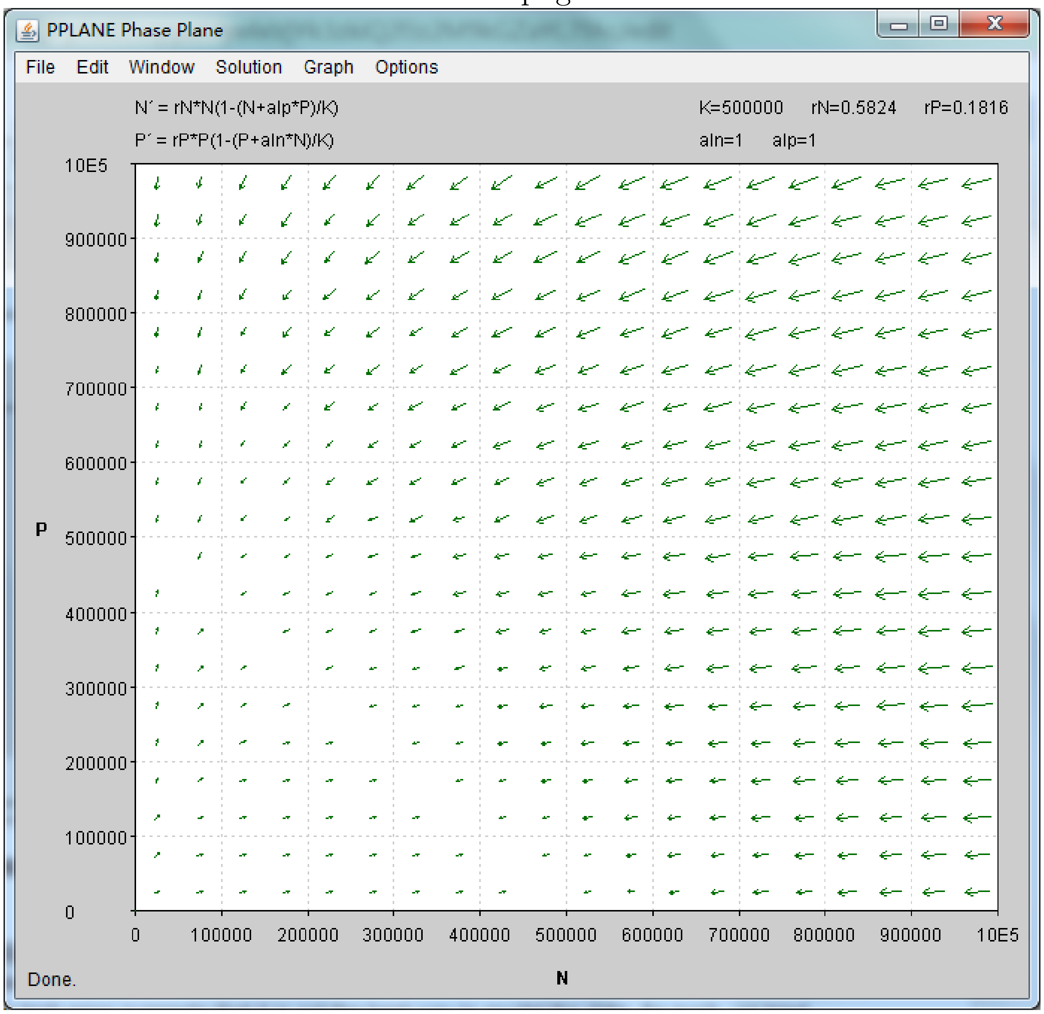


Errors:
Apple error: 9.5644e+08
Android error: 1.2064e+10
Combined error: 1.3020e+10

4.4.3 Stability Analysis

Below is the phase plane graph for the modified Lotka-Volterra Equations:

LV.png



This model has equilibrium solutions at the point N = P = 0 and at all points where N + P = K. This can easily be seen in the equations as well by setting $\frac{dN}{dt} = \frac{dP}{dt} = 0$.

Note that there are infinitely many stable equilibrium solutions corresponding to the line N + P = K, which means their maximum capacity depends on the initial value that we start with and the specific rates r_N and r_P. This can be easily seen with this model given the similarity between the equations for Android and Apple smartphone sales, as dividing the two equations gives the slope in the phase plane as

$$\frac{dN}{dP} = \frac{r_N N}{r_P P}$$

Which can be ly integrated and rearranged to give

$$\frac{ln(P)}{r_P} - \frac{ln(N)}{r_N} = c$$

Where c is a constant.

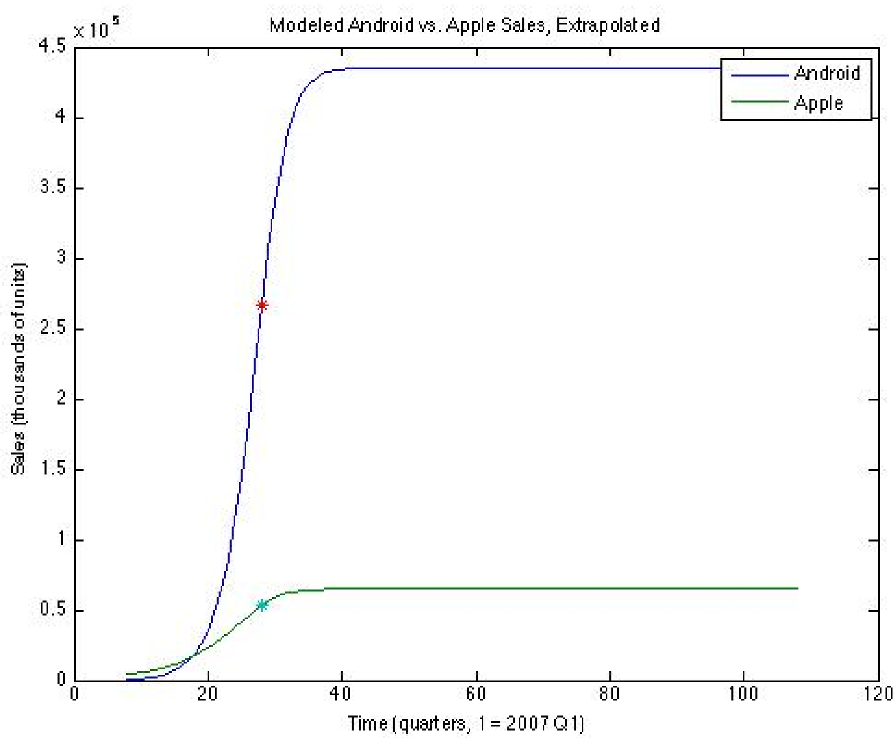
4.4.4 Extrapolating the Data

The data was also extrapolated by iterating the discrete form of the equation for a total of 100 time steps.

$$N(t + \Delta t) = N(1 + r_N \Delta t N(1 - \frac{N+P}{K})$$

$$P(t + \Delta t) = P(1 + r_P \Delta t P(1 - \frac{N+P}{K})$$

The marked points on each curve indicate where our data ends and the extrapolation begins.



4.4.5 Efficacy

From the graph we can see that this is a relatively good model, because it not only fits our initial data, but also predicts a reasonable trend in the future. It predicts that Android and Apple would grow until their combined sales reaches our estimated carrying capacity and then become stable. This model predicts the stable sales value for Android is 435,340 thousands of units while that for Apple is 64,660 thousands of units.

4.5 Modified Lotka-Volterra with Time-Varying Rates

While the equilibrium solutions of our modified model seem the most reasonable so far, the relatively high error suggests that it is not the best way to model the data. As such, we tried fitting a modified version of this model to the data that allowed the rates of increase for Apple and Android sales to vary annually. The equations are identical, except that we allowed the rates to be functions of time.

$$\frac{dN}{dt} = r_N(t)N\left(1 - \frac{N+P}{K}\right)$$

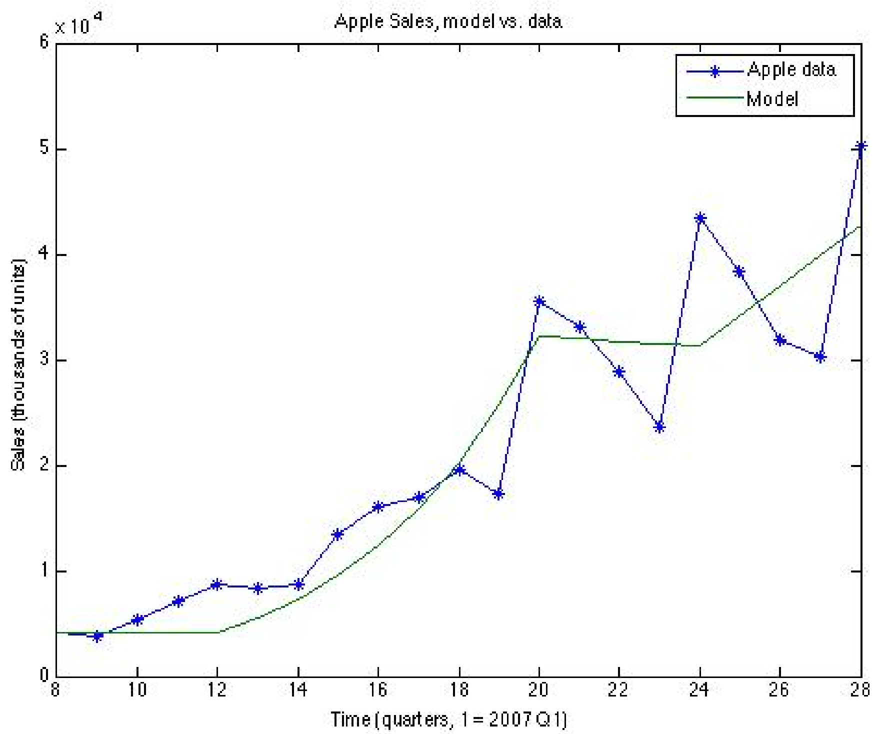
$$\frac{dP}{dt} = r_P(t)P\left(1 - \frac{P+N}{K}\right)$$

4.5.1 Additional Assumptions

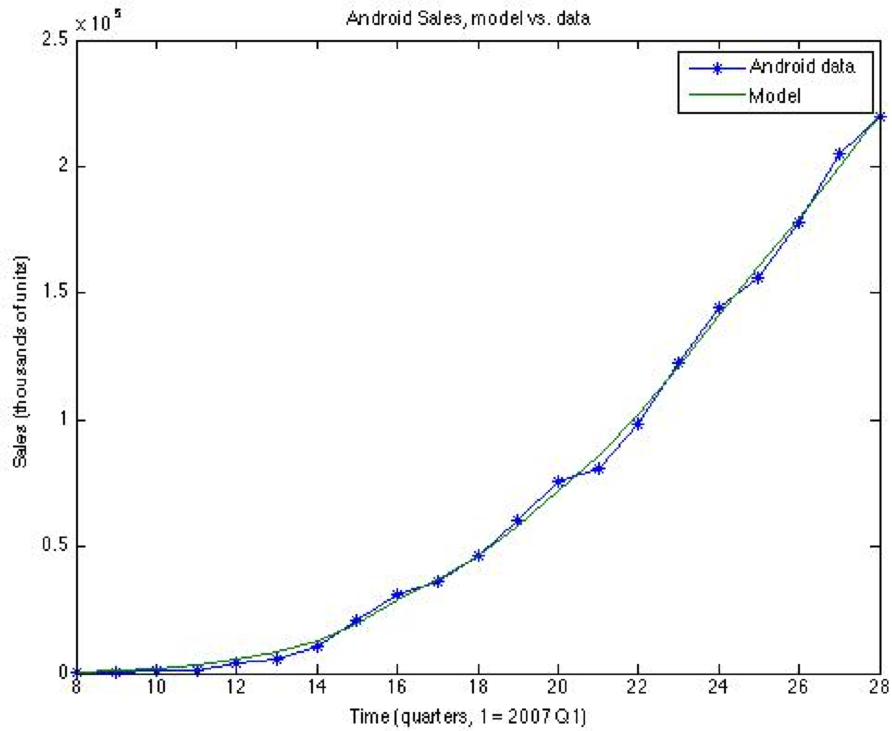
This model relies on all the same assumptions as the modified Lotka-Volterra except for the assumption that the rates of increase r_N and r_P are constant in time. Instead, we assumed that these rates could change once every four time steps (quarters). We initially planned to allow r_N and r_P to vary after every time step, but this proved to be too computationally intensive to optimize for, so we are approximating this behavior with annual rates.

4.5.2 Data Fitting

Below are the graphs we get when we calculate independent rate rN and rP for each year. Android:



Apple Sales, model vs. data
x 10^4
Apple data
Model
Sales (thousands of units)
Time (quarters, 1 = 2007 Q1)



Errors:
Apple error: 5.6180e+08
Android error: 1.4508e+08
Combined error: 7.0688e+08

4.5.3 Stability Analysis

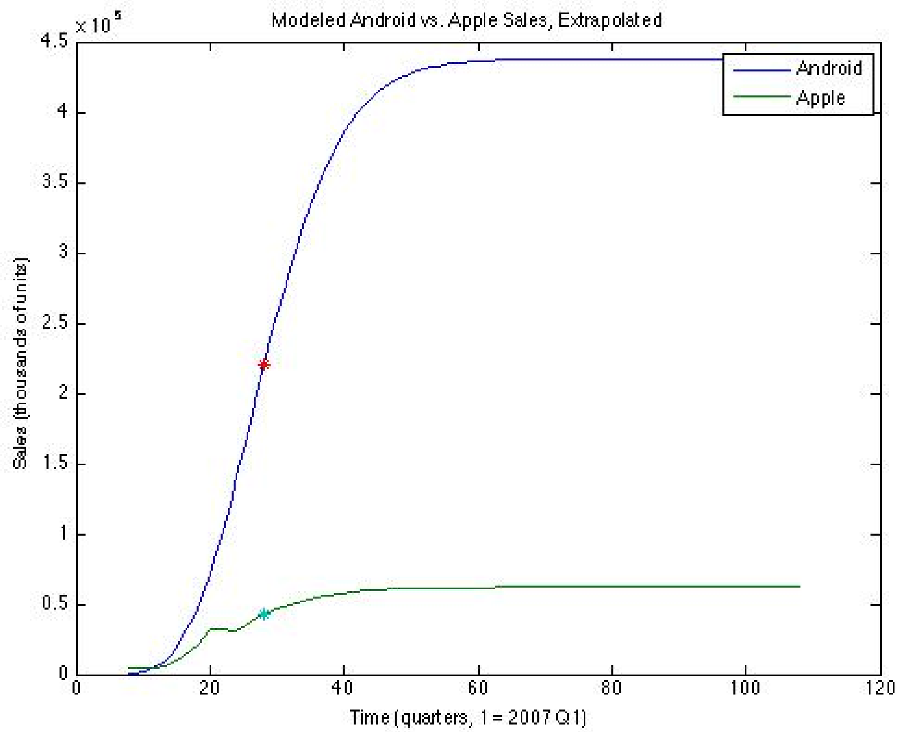
Since this model has nearly the same functional form as our previous model, we arrive at almost the same equation for curves in the phase plane, namely

$$\frac{ln(P)}{r_P(t)} - \frac{ln(N)}{r_N(t)} = c$$

However, the fact that we have left r_P and r_N as unknown functions of time leaves us at something of a dead end. Since we have no specific form for r_P and r_N, in order to examine the asymptotic behavior of this model we simply considered the average of the rates and relied on our previous stability analysis for the modified Lotka-Volterra model.

4.5.4 Extrapolating the Data

From this diagram, it shows that these two trends are really closed to the model we are trying to get from original Apple and Android data. Without extraneous factors, Andriod and Apple sales will keep booming and in 25 years and reach a stable capacity, which is extremely similar (as expected) to that predicted in the original modified Lotka-Volterra model.



4.5.5 Efficacy

Comparing the errors obtained by using this model or by visual inspection, we see that, even with rates changing only once a year, this model does significantly better than any we tried previously at matching the data. Additionally, its asymptotic behavior is tied for the most reasonable with the previous attempt at a modified Lotka-Volterra model.

5 Summary and Conclusion

Error values for different models (in thousands of units squared):

Model	Apple Error	Android Error	Combined Error
Logistic	7.2969e+08	1.4764e+09	2.2061e+09
LV Predation	1.5863e+09	5.1929e+10	5.3515e+10
Competitive LV, Indep r	5.8270e+08	8.9723e+08	1.4799e+09
Modified LV	9.5644e+08	1.2064e+10	1.3020e+10
Modified LV, r(t)	5.6180e+08	1.4508e+08	7.0688e+8

Examining the errors in the above table, we see that few of the models perform better than our simple logistic fit. Unsurprisingly, the best fit to the data occurs when we allow the rate to vary with time, as this corresponds to the most parameters that can be fit to the data. What is surprising is that the 2 parameter modified Lotka-Volterra fit results in a worse overall error than 2 distinct one parameter Logistic fits. This suggests that the smartphone sales are better modeled by treating the two companies as more independent than not, as modeling their behavior as more interdependent does not improve how well the data can be fit.

Predicted Sales after 25+ years (in thousands of units):

Model	Apple Sales	Android Sales	Combined Error
Logistic	5.0000e+05	5.0000e+05	1.0000e+06
LV Predation	∞	∞	∞
Competitive LV, Indep r	0.0121	5.0000e+05	5.0000e+05
Modified LV	6.4660e+04	4.3534e+05	5.0000e+05
Modified LV, r(t)	6.1931e+04	4.3807e+05	5.0000e+05

However, looking at the future sales predicted by these models suggests that many of those that had better overall errors are poor models when it comes to explaining the data. Only the last 3 models we attempted exhibited asymptotic behavior that satisfied the condition that total phone sales would approach the carrying capacity. Of these, only the modified Lotka-Volterra equations predict sales that seem reasonable, as it does not appear likely that Android will drive Apple phone sales into the ground in the next couple of years.

The model that performs the best is our modified Lotka-Volterra model that allows the rate to change with time. This suggests that our initial goal of treating smartphone sales by different companies as competing populations is infeasible, as this model assumes that the rates of increase of each population is more or less constant and ignores the effects of any factors that would change it. As such, we conclude that the various specific economic factors that we ignored when constructing our models are essential to developing an accurate understanding of the nuances of the smartphone market.

6 references

"Gartner Smart Phone Marketshare 2011 Q3". Gartner, Inc.

"Gartner Smart Phone Marketshare 2011 Q4". Gartner, Inc.

"Gartner Smart Phone Marketshare 2012 Q1". Gartner, Inc.

"Gartner Smart Phone Marketshare 2012 Q2". Gartner, Inc.

"Gartner Smart Phone Marketshare 2012 Q3". Gartner, Inc.

"Gartner Smart Phone Marketshare 2013 Q1". Gartner, Inc.

"Gartner Smart Phone Marketshare 2013 Q2". Gartner, Inc.

"Gartner Smart Phone Marketshare 2013 Q3". Gartner, Inc.

."Gartner Smart Phone Marketshare 2013 Q4". Gartner, Inc.

"Gartner Smart Phone Marketshare 2011 Q2". Gartner, Inc.

"Gartner Smart Phone Marketshare 2011 Q1". Gartner, Inc.

"Gartner Smart Phone Marketshare 2010 Q4". Gartner, Inc..

"Gartner Smart Phone Marketshare 2010 Q3". Gartner, Inc.

"Gartner Smart Phone Marketshare 2010 Q2". Gartner, Inc.

"Gartner Smart Phone Marketshare 2010 Q1". Gartner, Inc.

Rice University, Dept. Of Mathematics, : http://math.rice.edu/ dfield/dfpp.html