Introduction to Quantitative Biology

K. Sriram

Lecture-1

Course book


Mathematical Models in Biology
An Introduction
ELIZABETH S. ALLMAN
JOHN A. RHODES
CAMBRIDGE
CAMBRIDGE
more information - www.cambridge.org/9780521819800

Tentative Topics

No	Topics	What quantitative model part will be covered?	MATLAB/Python programming Required?
1.	Population biology (discrete dynamical models)	Difference equation	Yes
2.	Age structured models	Matrix algebra	Yes
3.	Prey-predator models (Continuous dynamical models)	Differential equations	Yes
4.	Molecular evolution	Probability	Yes
5.	Phylogenetics	Markov models	Yes
6.	Genetics	Probability	Yes
7.	Curve Fitting and Biological Modeling	Linear algebra	Yes

All will be covered at appropriate introductory level

Evaluation

No	Exam	Weightage
1.	Quizzes-I & II	20
2.	Mid-Sem	20
3.	End-Sem	30
4.	Assignments	30
5.	Home work problems (from prescribed book)	0

Office hours: Monday 2-3 PM @office A-308, R&D block.

TA's : Will let you know next week.

Slides and other materials: Will be uploaded in google classroom.

Quantitative models

- Quantitative models are the caricature of the real representation of events happening around us.

- When to model?
- What to model?
- How good is the model?
- How to validate the model?
- How to use the model to make predictions?

- One example of quantitative model is from population biology.

Quantitative biology of population growth

- A population is the group of plants, animals, or other organisms, all of the same species that live together and reproduce.

- Variable P is used to indicate the size of the population.

- Since population changes with time ‘t’, we indicate the time with subscript P_t .

- We use $t = 0$, the starting time as a standard convention.

Example

- Take the census of spider for this year and next year.
- P_0 is the this time of year where the spider population is say 500.
- P_1 is the population of spider the consecutive year which is say 800 (increased from 500).
- **<u>Aim</u>**: In general we are interested in predicting the future population from the present population.
- i.e., Future population of P_{t+1} from present population P_t

Quantifying changes in the population size

- All the changes in the population size can be classified in four types.

- Increase in population due to birth (B)
- Decrease in population due to death (D)
- Increase in population due to immigration (I)
- Decrease in population due to emigration (E)

- These four combinations will change the population size.

Mathematics of population Change: Malthusian model

- The mathematical representation for the changes in population is given by

$$P_{t+1} = P_t + B - D + I - E$$

$$P_{t+1} - P_t = B - D + I - E$$

$$\boxed{\Delta P \text{ (change in the population)} = B - D + I - E}$$

- Assume that it's a closed population; i.e., no immigration or emigration.

- Then the change in the population becomes

$$\Delta P = B - D$$

Birth and death in terms of population 'P'

- Birth and death rates depends on present population size ' P_t '

- Let 'f' be the instantaneous birth rate so that

$$B = f\,P_t$$

- Similarly ' d ' is the instantaneous death rate so that

$$D = d\,P_t$$

- Inserting these terms in the change in population equation (ΔP), we get

$$\Delta P = B - D = f\,P_t - d\,P_t = (f - d)\,P_t$$

$$P_{t+1} - P_t = (f - d)\,P_t$$

Difference equation

- The above equation can be written in the form

$$P_{t+1} - P_t = (f - d)\, P_t \quad \text{Since } \Delta P = P_{t+1} - P_t$$

$$P_{t+1} = P_t + (f - d)\, P_t$$

$$P_{t+1} = (1 + f - d)\, P_t$$

Put $\lambda = (1 + f - d)$ and this leads to

$$\boxed{\begin{array}{c} P_{t+1} = \lambda\, P_t \\ \lambda = \text{finite growth rate} \end{array}}$$

Example and extension

- For the values f = 0.1, d = .03, and the initial value of population $P_0 = 500$, we get

$$\lambda = (1 + f - d) = (1 + 0.1 - 0.03)$$

$$P_{t+1} = 1.07P_t\ , \quad \text{with initial condition } P_0 = 500$$

Day (t)	Population (P_t)
0	500
1	(1.07) 500 = 535
2	$(1.07)^2$ 500
3	$(1.07)^3$ 500
:	:

$P_0 = 500$

$P_1 = 1.07P_0$

$P_2 = 1.07P_1 = (1.07)^2P_0$

$P_3 = 1.07P_2 = (1.07)(1.07)^2P_0$

:

$\boxed{P_{t+1} = (1.07)^t P_t}$ **(DIFFERENCE EQUATION)**

Also called linear model that produces exponential growth

What is a difference equation?

- A difference equation is a formula expressing values of some quantity *Q* in terms of previous values of *Q*

- Thus, if $F(x)$ is any function, then

  $Q_{t+1} = F(Q_t)$ (In the previous example, $F(x) = \lambda x$)

  is called a difference equation

- In this course we will address two main issues:

  1) How do we find an appropriate difference equation to model a situation?

  2) How do we understand the behavior of the difference equation model once we have found it?

Flaws in the earlier model

- The Malthusian model predicts that population growth will be exponential.

- Exponential functions grow quickly and without bound.

- The main flaw in the assumption is the growth and death rates for our population are the same regardless of the size of the population.

- So when population gets large, it might be more reasonable to expect a higher death rate and a lower fecundity.

- We need to somehow modify our model so that the growth rate depends on the size of the population; that is, the growth rate should be *density dependent*.

Creating a nonlinear model

- We want to create a model that satisfies the following conditions

- For small values of *P, the per-capita growth rate* (ΔP/P) *should be large, since it's a* small population with lots of resources available in its environment to support further growth.

- For large values of *P, however, per-capita growth* should be much smaller, as individuals compete for both food and space.

- For even larger values of *P, the per-capita growth rate should be negative, since* that would mean the population will decline.

Graph expected for ΔP/P vs. P

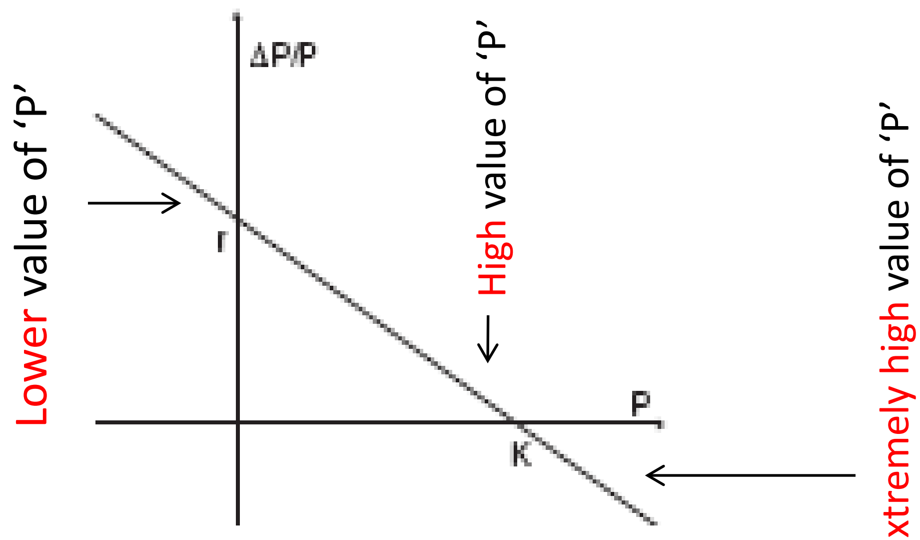
Extra loss term added.
So it becomes nonlinear in P (i.e., P^2).

$$\frac{\Delta P}{P} = r\left(1 - \frac{P}{K}\right)$$

Carrying capacity of the population

Compare with earlier linear equation:

$$\Delta P = B - D = f\,P_t - d\,P_t$$
$$= (f - d)\,P_t$$
$$\Delta P/\,P_t = (f -- d) = r$$



HW: Problems from book

- Problems 1.1.1, 1.1.2, 1.1.6, and 1.1.7

Lecture-2

Cobweb diagrams, equilibrium and stability

Logistic model

$$\frac{\Delta P}{P} = r\,(1 - \frac{P}{K})$$

There are two parameters in the model; *r* and *K (How to estimate?)*

Limiting Case-1: P < < K, P/K is small, so $\Delta P/P \approx r > 0$

$$P_{t+1} - P_t \approx r\,P_t \;\text{ or }\; P_{t+1} \approx P_t + r\,P_t \approx (1+r)\,P_t$$

Limiting Case-2: P > > K P/K is very large & >1, so $\Delta P/P \approx -r < 0$

- K is the carrying capacity, that represents the maximum number of individuals that can be supported over a long period.
- *r (= f - d) simply reflects the way* the population would grow or decline in the absence of density-dependent effects – when the population is far below the carrying capacity.

Different representation of ΔP *or* P_{t+1}

$$\boxed{\frac{\Delta P}{P} = r\left(1 - \frac{P}{K}\right)}$$

- (i) $\Delta P = r\,P(1 - P/K)$
- (ii) $\Delta P = sP(K - P)$
- (iii) $\Delta P = t\,P - uP^2$
- (iv) $P_{t+1} = vP_t - wP_t^{\,2}$

- *Exercise*: Write each of the following in all four of these forms. (problem 1.2.5)

  (i) $P_{t+1} = 2.5P_t - 0.2P_t^{\,2}$.

  (ii) $P_{t+1} = P_t + 0.2\,P_t\,(10 - P_t)$.

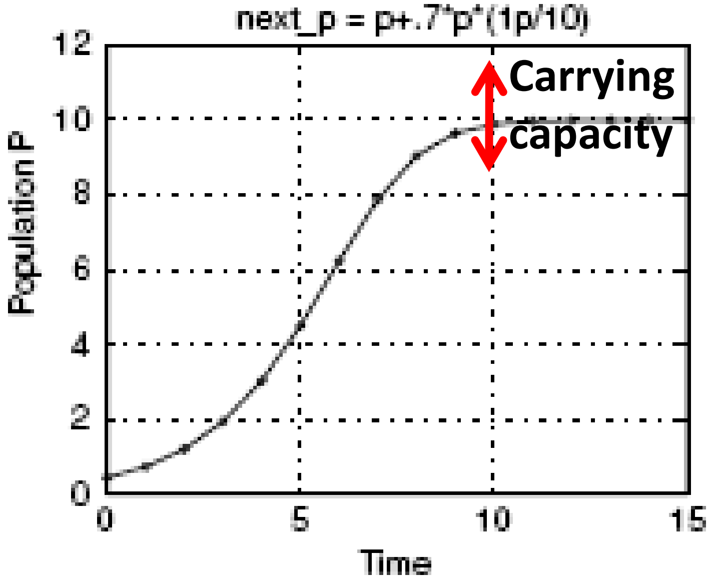
Iterating the model

- Consider the model

$$P_{t+1} = P_t\,(1 + .7(1 - P_t/10)) \text{ and } P_0 = 0.4346$$

Table 1.5. *Population Values from a Nonlinear Model*

t	0	1	2	3	4	5	6
P_t	.4346	.7256	1.1967	1.9341	3.0262	4.5034	6.2362
t	7	8	9	10	11	12	...
P_t	7.8792	9.0489	9.6514	9.8869	9.9652	9.9895	...



- **Nonlinear model, very realistic. Seen in lab experiments.**
- **Curve is sigmoidal.**
- **Reaches saturation close to carrying capacity**

Cobweb diagrams for discrete systems Graphical method

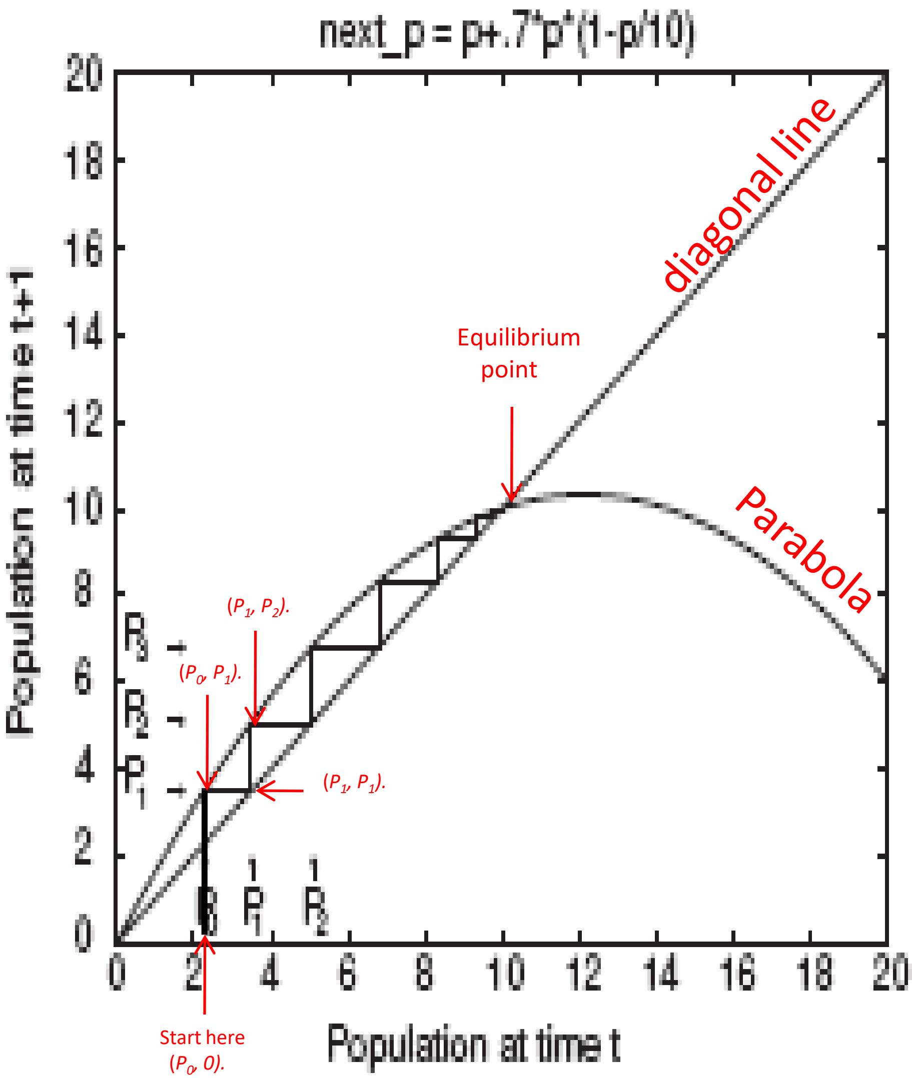
- Cobwebbing is the basic graphical technique for understanding a model such as the discrete logistic equation.

- One can understand the complex dynamics, including the equilibrium points.

- Let us start with the simple logistic equation of the form

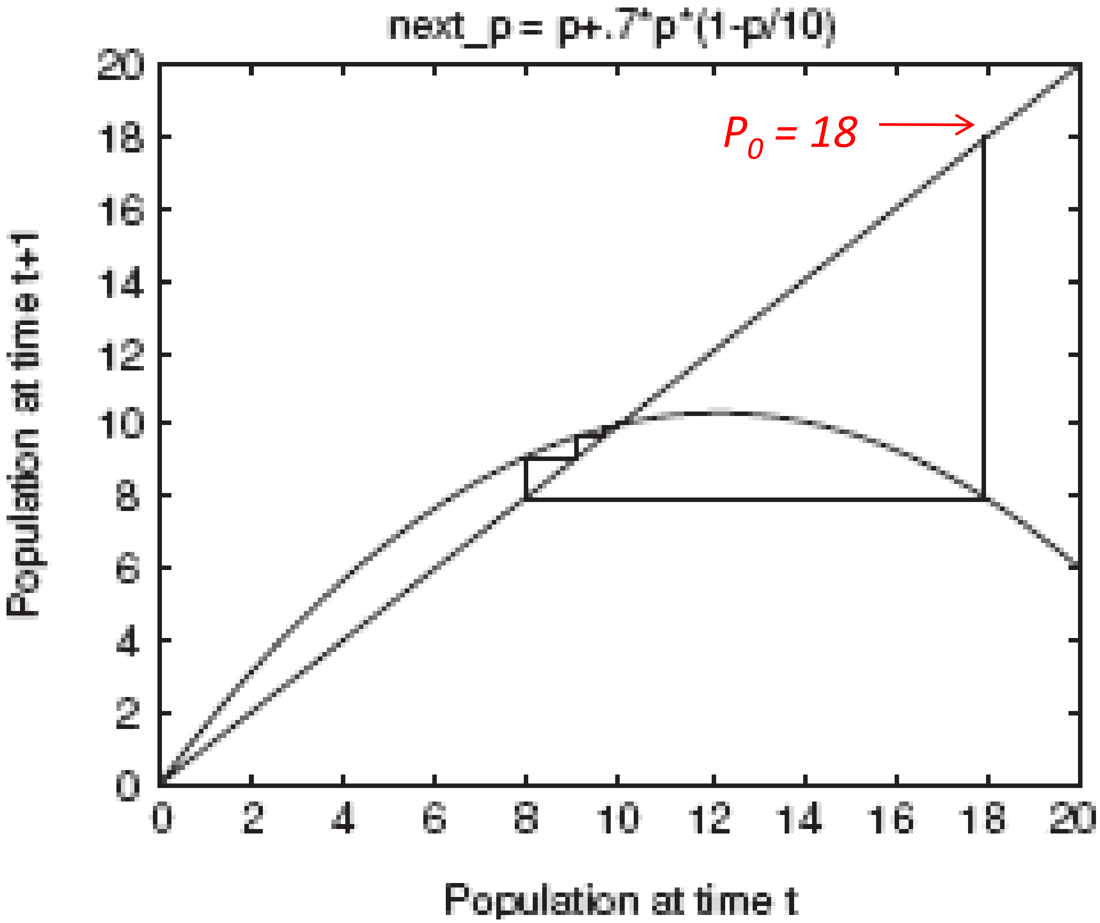
$$P_0 = 2.3, \; P_{t+1} = P_t\,(1 + .7(1 - P_t/10))$$

Steps to graph the cobweb plot



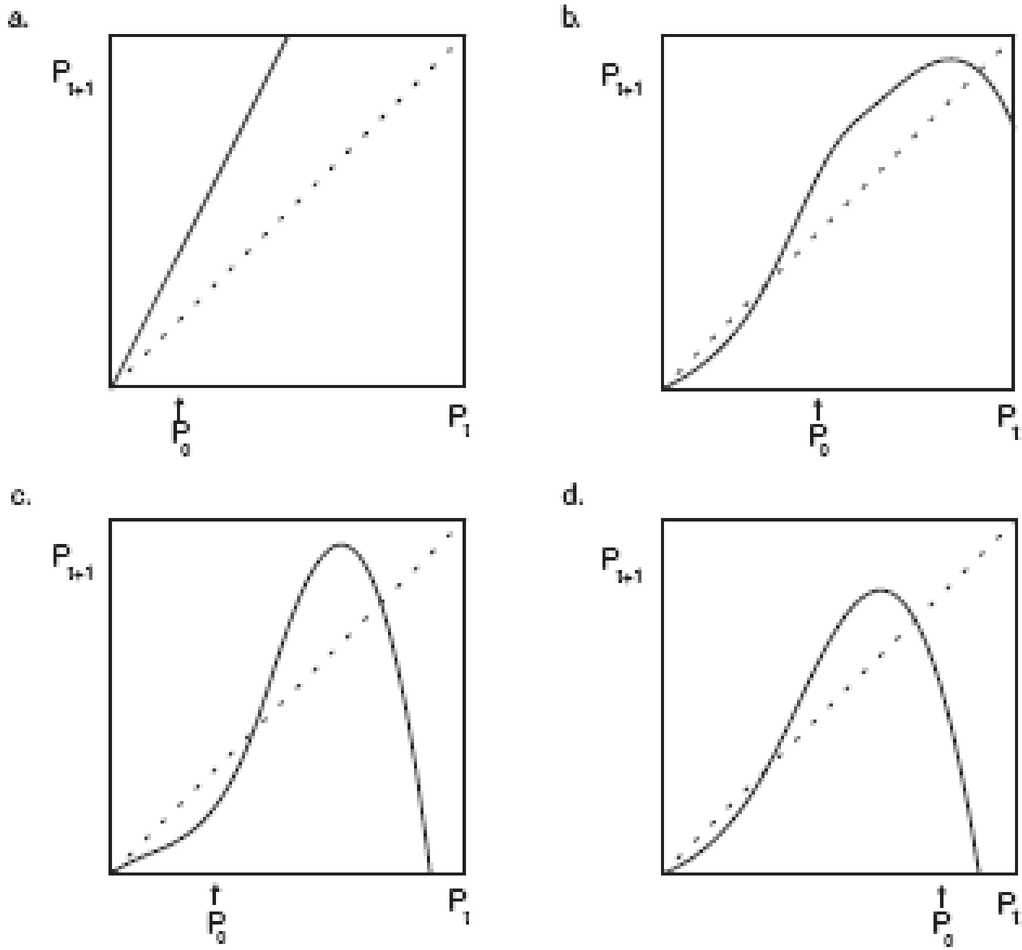
- It's a plot of P_t Vs P_{t+1}
- First draw a diagonal or 45^0 straight line which $P_{t+1} = P_t$.
- Draw the graph of P_{t+1} Which gives parabola.
- Since the population begins at $P_0 = 2.3$, we mark that on the graph's horizontal axis. i.e., $(P_0,0)$.
- Now, to find P1, we just move vertically upward to the graph of the parabola to find the point (P_0, P_1).
- To find P_2 next, we hit the diagonal line, we will be at (P_1, P_1) and then move vertically back to the parabola to find the point (P_1, P_2).
- Repeat the same process.

Start with different initial condition



If the population becomes negative, then we should interpret that as extinction.

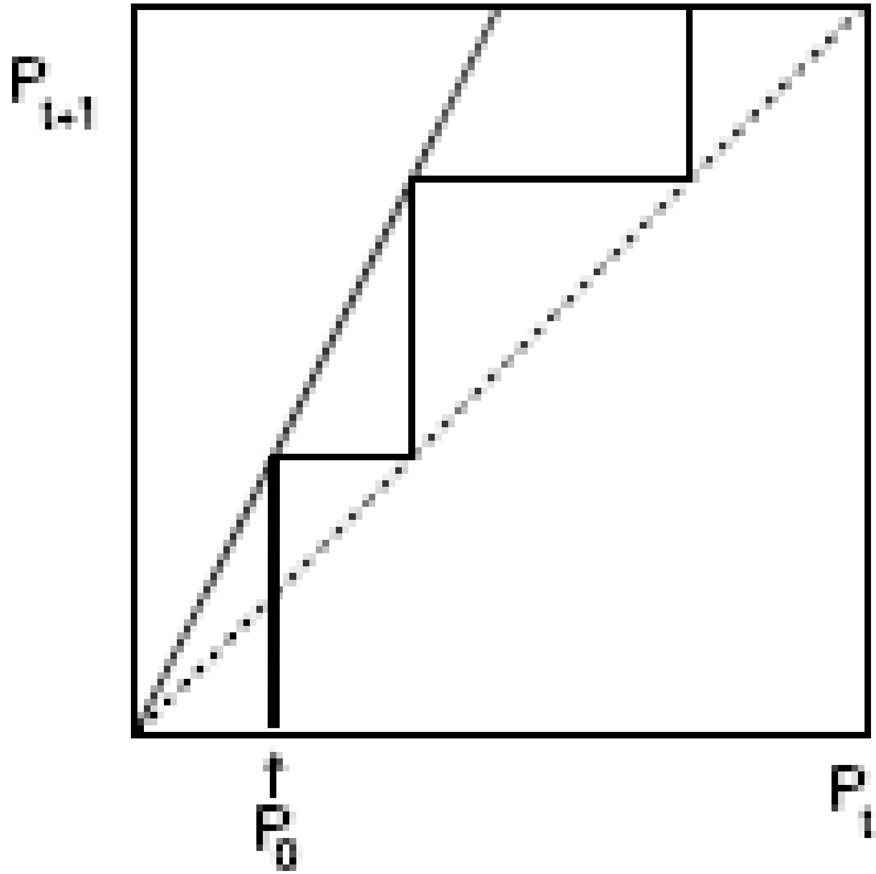
In class exercise: Problem 1.2.9



Give a formula for the graph appearing in part (a) of the Figure. What is the name of this population model?

Solution-a

a.

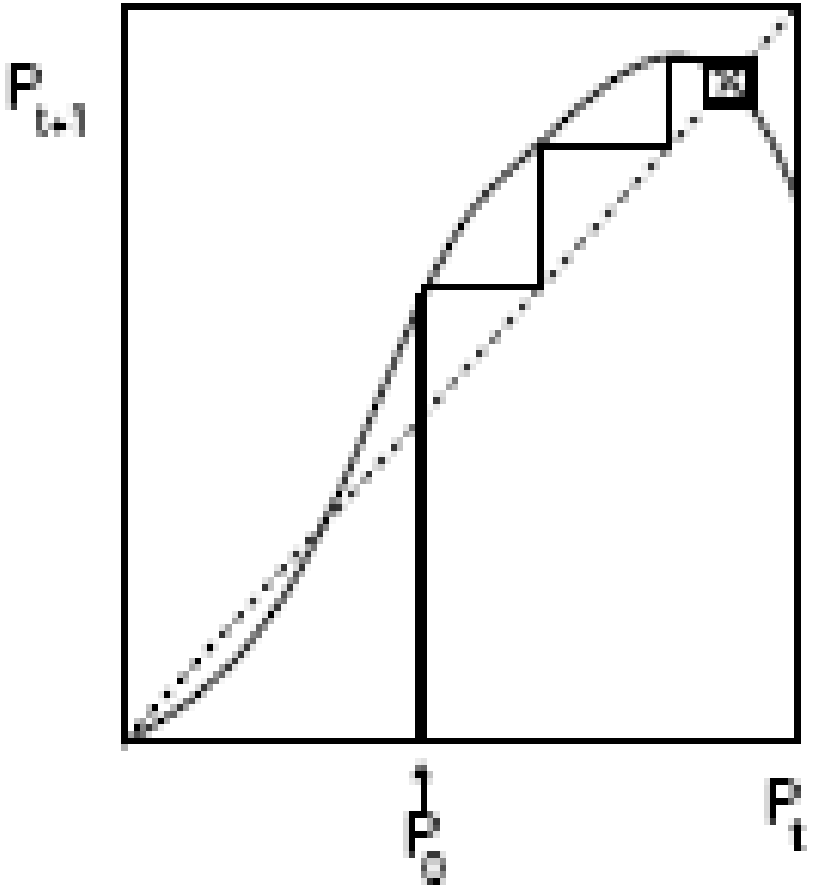


- Since the graph appears to be a straight line through the origin with slope 2, the model is

$$P_{t+1} = 2P_t$$

- *This is an exponential growth model*

Solution-b

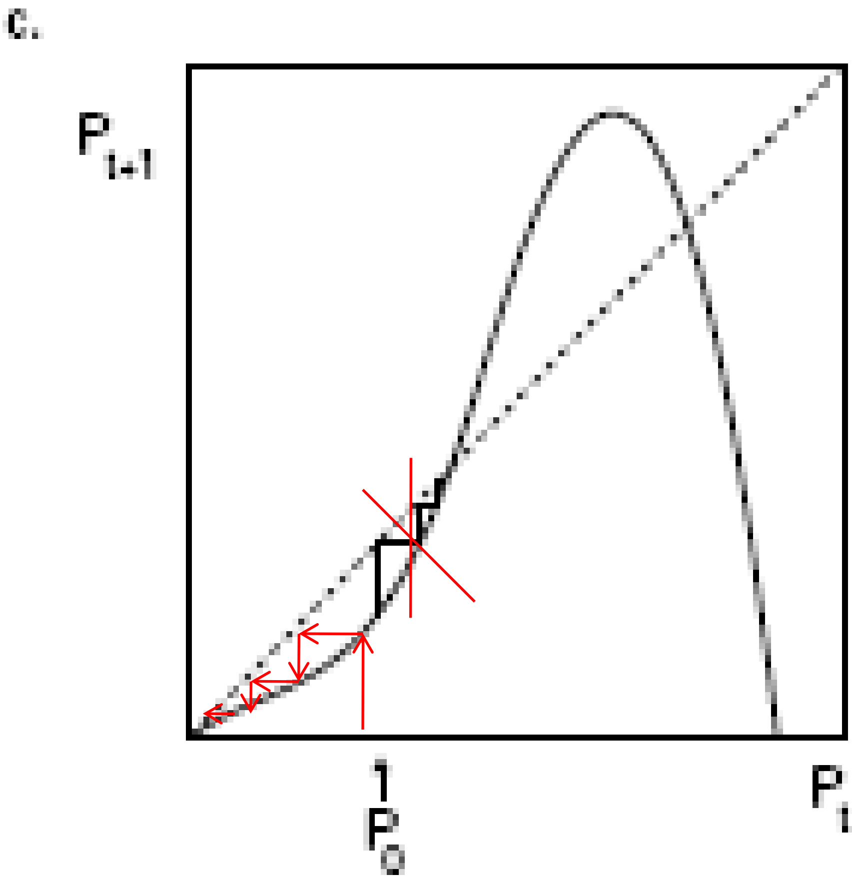


From graph,
How many equilibrium points are there?

Can you guess what can be the function $F(P_t)$? i.e.,

$$P_{t+1} = F(P_t)$$

Solution-c



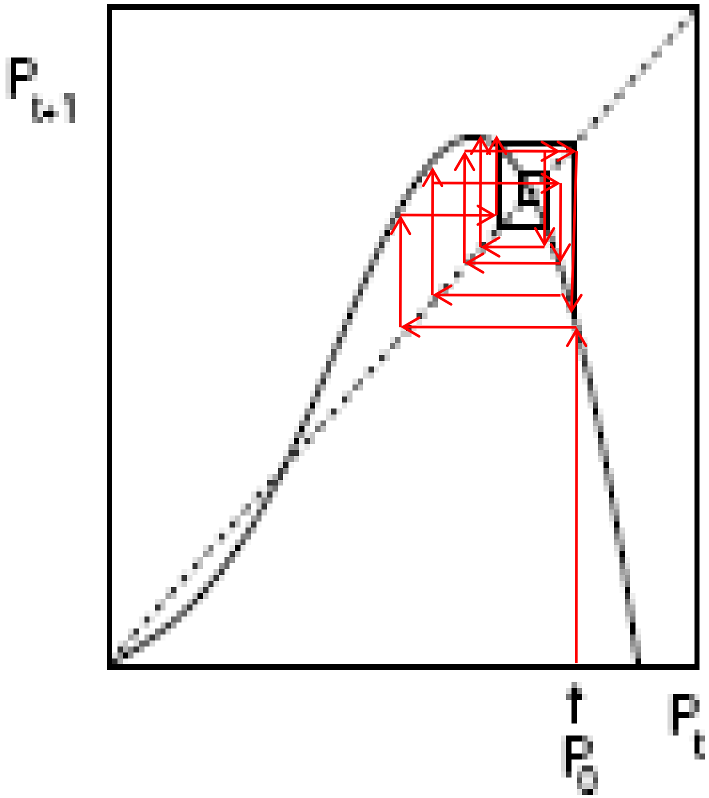
From graph,
How many equilibrium points are there?

Can you guess what can be the function $F(P_t)$? i.e.,

$$P_{t+1} = F(P_t)$$

Solution-d

d.



From graph,
How many equilibrium points are there?

Can you guess what can be the function $F(P_t)$? i.e.,

$$P_{t+1} = F(P_t)$$

Analyzing Nonlinear Models Transients, equilibrium, and stability

- Take the discrete logistic model

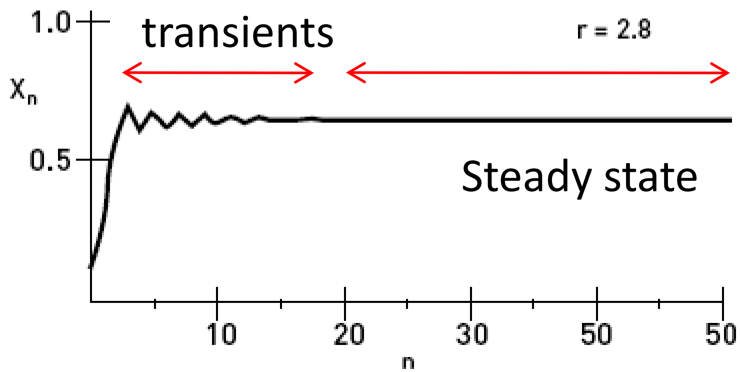
$$P_{t+1} = P_t\,(1 + .7(1 - P_t\,/10))$$

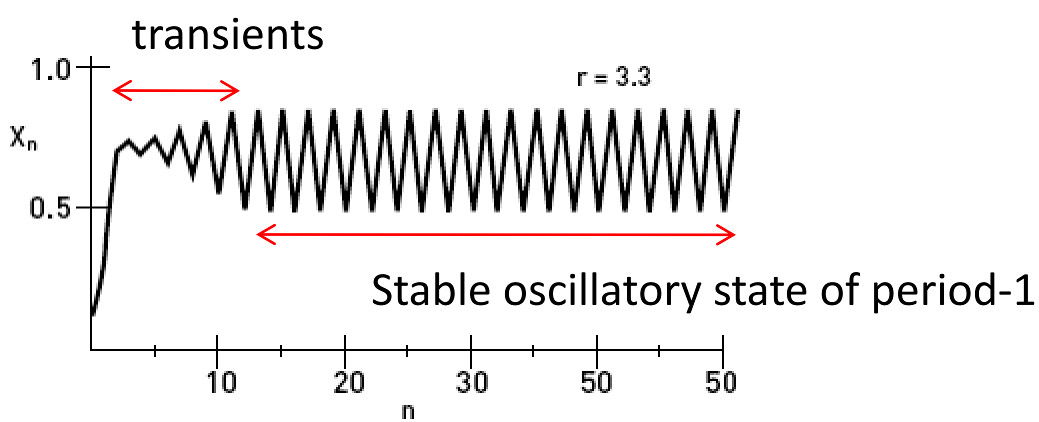
- This moves towards equilibrium point 10.

- However, starting from initial values P_0, the first few iterations of the model produce relatively large changes in Pt as it moves toward 10.

- This early behavior is thus called transient, because it is ultimately replaced with a different sort of behavior.

Examples of transient from time series i.e., X_n vs n for $X_{n+1} = rX_n(1-X_n)$





Lecture-3

Equilibria and Linearization around equilibria

Equilibrium

- **Definition**. An *equilibrium value for a model* $P_{t+1} = F(P_t)$ *is a value* P^* such that $P^* = F(P^*)$. *Equivalently, for a model* $\Delta P = G(P_t)$, *it is a value* P^* *such that* $G(P^*) = 0$.

- **Problem**: Finding equilibrium values for the model

$$P_{t+1} = P_t\,(1 + .7(1 - P_t/10))$$

- **Solution**: Let P^* *be the equilibrium point such that*

$$P^* = P^*(1 + .7(1 - P^*/10))$$

- There are precisely two equilibrium values

$$P^* = 0 \text{ or } P^* = 10$$

Linearization around the equilibrium point

- Stability is determined close to an equilibrium point.
- *We look at dynamics close to equilibrium point P^**
- we consider a population $P_t = P^* + p_t$ *, where* p_t *is a very* small number that tells us how far the population is from equilibrium P^*.
- (**NOTE**: P_t is different from p_t)
- P_t *is the perturbation from equilibrium.*
- We then compute $P_{t+1} = P^* + p_{t+1}$ *and use it to find* p_{t+1}.
- If p_{t+1} *is* bigger than p_t *in absolute value, then we know that* p_{t+1} *has moved away* from P^*.
- If p_{t+1} *is* smaller than p_t *in absolute value, then we know that* p_{t+1} *has moved away* from P^*.

Example to determine local stability of equilibria

- Consider again the model

$$P_{t+1} = P_t\,(1 + .7(1 - P_t/10))$$

- *There are two* equilibria; $P^* = 0$ *and* 10

- Consider the equilibria $P^* = 10$;

- Substituting $P_t = 10 + p_t$ *and* and $P_{t+1} = 10 + p_{t+1}$ *into the equation for the model yields:*

$$10 + p_{t+1} = (10 + p_t\,)(1 + .7(1 - (10 + p_t\,)/10))$$

Example to determine stability of equilibria

After all algebra, the final equation reads as

$$p_{t+1} = 0.3p_t - .07p^2_{\mathrm{t}} .$$

$p_{\mathrm{t+1}} \approx 0.3p_{\mathrm{t}}$ (Note: This is perturbation p_t)

- Equilibrium is compressed by a factor of about 0.*3 with each time step.*
- Small perturbations from the equilibrium therefore shrink.
- *Therefore P* = 10 is* indeed stable

What is Linearization?

- The process performed in this example is called *linearization of the model* at the equilibrium.

- We first focus attention near the equilibrium by our substitution $P_t = P^* + p_t$, *and then ignore the terms of degree greater* than 1 in p_t.

- Linear models, as we have seen, are easy to understand, because they produce either exponential growth or decay.

Linearization around P* =0 of the model

- Linearization at P*$= 0$ *yields*

$$p_{t+1} = 1.7p_t$$

- Therefore, perturbations from this equilibrium grow over time, so $P^* = 0$ *is* unstable.

- In general, when the stretching factor is greater than 1 in absolute value, the equilibrium is unstable.

- When it's less than 1 in absolute value, the equilibrium is stable.

Calculus to understand stability

Take the ratio of the linear system P_{t+1} / P_t

$$\frac{p_{t+1}}{p_t} = \frac{P_{t+1} - P^*}{P_t - P^*} = \frac{F(P_t) - P^*}{P_t - P^*} = \frac{F(P_t) - F(P^*)}{P_t - P^*},$$

$$\lim_{P_t \to P^*} \frac{F(P_t) - F(P^*)}{P_t - P^*}. \quad = |F'(P^*)| = dF/dP \text{ (at } P = P^*)$$

Theorem on stability. If a model $P_{t+1} = F(P_t)$ has equilibrium P∗, then |F'(P*)| > 1 implies P* is unstable, while |F'(P*)| < 1 implies P∗ is stable. If |F'(P*)| = 1, then this information is not enough to determine stability.

Example

- Example: Using $P_{t+1} = P_t\,(1 + .7(1 - P_t/10))$
- so $F(P) = P(1 + .7(1 - P/10))$,
- Compute $F'(P) = (1 + .7(1 - P/10)) + P(.7)(-1/10)$.
- $F'(10) = 1 - .7 = 0.3$, and $P^* = 10$ is stable
- H.W. Check the stability for other fixed point, $P^*= 0$.

Variations on the Logistic Model

- The model $P_{t+1} = P_t\,(1 + .7(1 - P_t\,/10))$ is unrealistic.

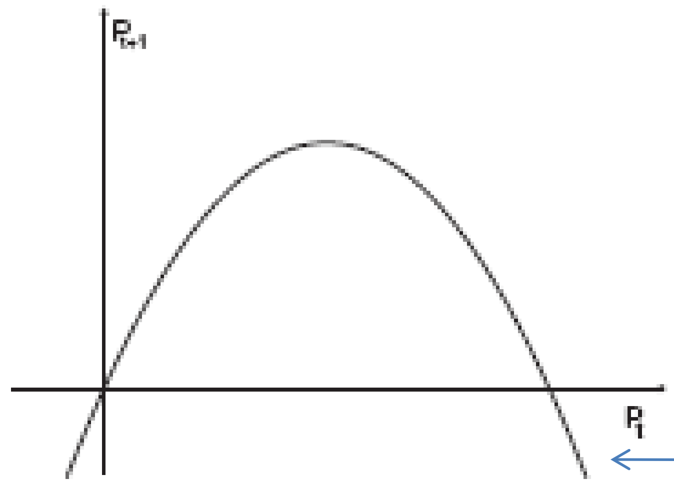


Figure 1.9. Model with unrealistic $P_{t+1} < 0$ for large P_t.

Becomes negative

- We interpret the negative population as extinct.

Realistic models

- Perhaps a more reasonable model would have large values of P_t *produce* very small (but still positive) values of P_{t+1}.
- *Thus, a population well over* the carrying capacity might immediately crash to very low levels, but at least some of the population would survive

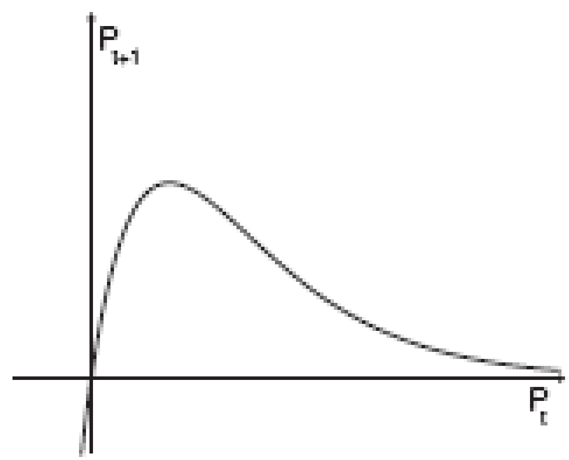


Figure 1.10. New model with $P_{t+1} > 0$.

Ricker and other models

- ***Model-1:*** *Ricker model:* $P_{t+1} = P_t e^{r(1-Pt/K)}$
- ***Model-2:*** $P_{t+1} = \lambda P_t/(1 + aP_t)\beta.$

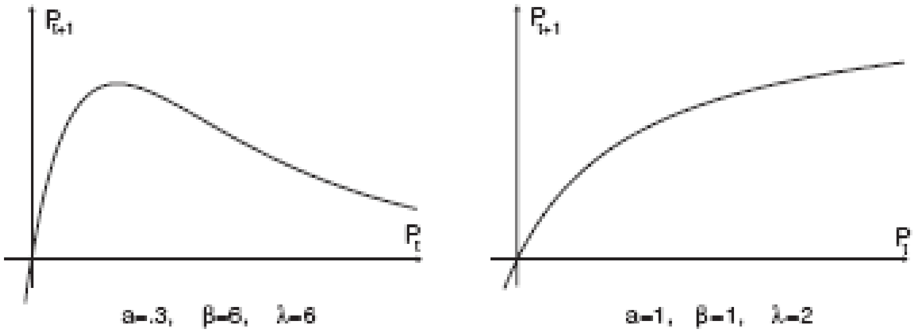


Figure 1.13. Two models of the form $P_{t+1} = \frac{\lambda P_t}{(1+aP_t)^\beta}$.

- **H.W**: (i) Compute the equilibria for both model and how many equilibrium points are there?
- (ii) Draw the cobweb plot and identify the equilibrium points graphically
- (iii) Draw the plot of per-capita growth rate.

Lecture-4

Oscillations, Chaos and In-class problem solving

Oscillations, bifurcations and chaos

- Consider the logistic model $\Delta P = r\,P(1 - P/K)$.

- Setting $K = 1$, for any value of r the logistic model has two equilibria, 0 and 1.

- The "stretching factor" at $P^* = 0$ is $1 + r$; i.e,

$$p_{t+1} \approx (1 + r\,)p_t \quad \text{(show it!)}$$

- $P^* = 0$, then is always an unstable equilibrium for

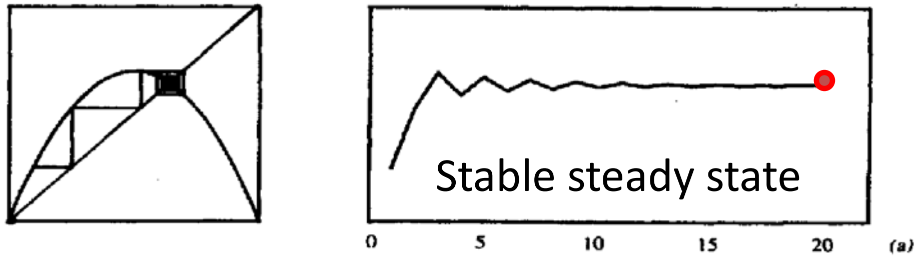
  $r > 0$.

- The "stretching factor" at $P^* = 1$ is $1 - r$ (show it!).

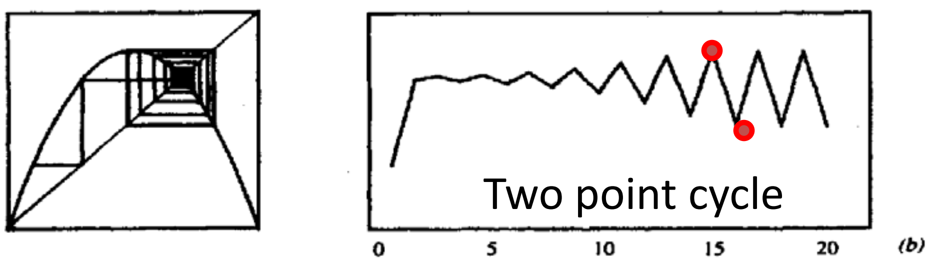
How oscillation arise?

- Consider the equilibrium point $P^* = 1$
- First, when $0 < r \leq 1$, then $0 \leq 1 - r < 1$, so the equilibrium is stable.
- The formula $p_{t+1} \approx (1 - r)p_t$ shows that the population simply moves toward equilibrium without ever overshooting it.
- When r is increased so that $1 < r < 2$, then $-1 < 1 - r < 0$ and the equilibrium is still stable.
- However, the sign of p_t will alternate between positive and negative as t increases.
- oscillatory behavior is observed above and below the equilibrium as our perturbation from equilibrium alternates in sign.
- The population therefore approaches the equilibrium as a damped oscillation.

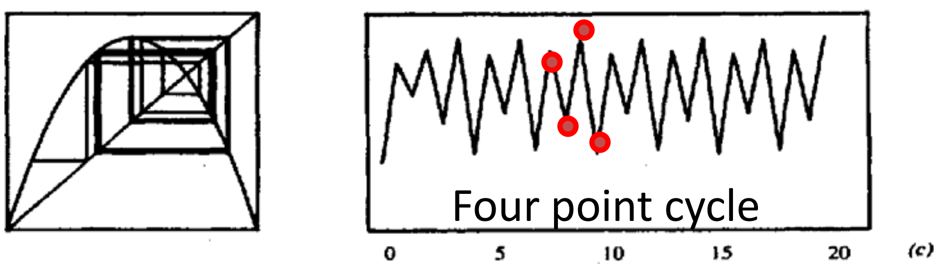
r = 1.8



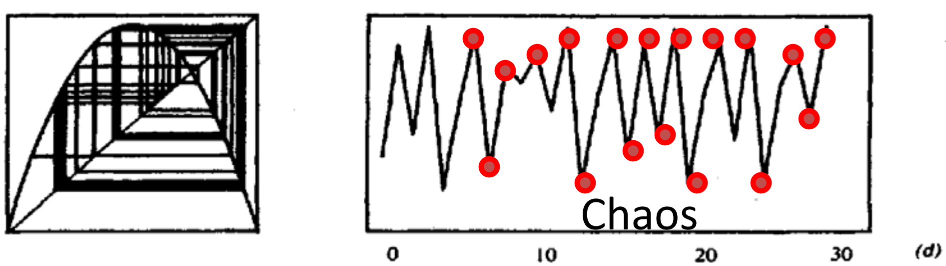
r = 2.2



r =2.55



r = 2.9



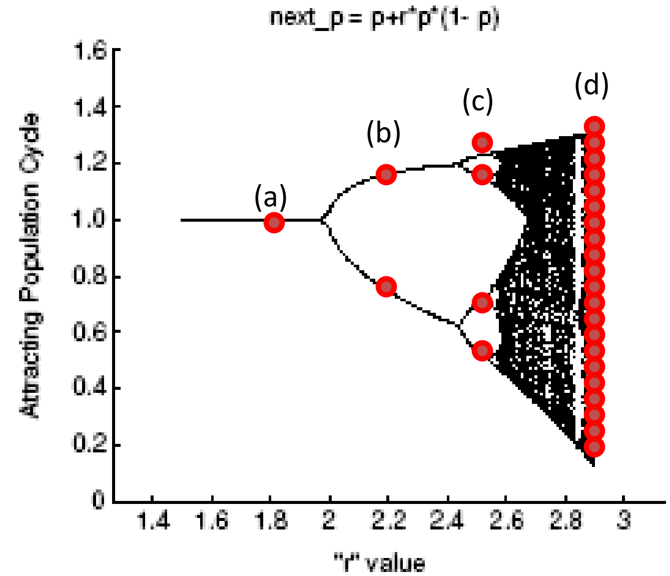


Figure 1.6. Bifurcation diagram for the logistic model.

- When *r is* increased beyond 2, we find |1 − *r| exceeds 1 and therefore the stable equilibrium* at *P∗ = 1 becomes unstable.*

- *Thus, a dramatic qualitative behavior* change occurs as the parameter is increased across the value 2.

The 2-cycle becoming a 4-cycle. Further increases in *r produce an 8-cycle,* then a 16-cycle, and so on, and finally to chaos.

Summary

- Two population models are studied.

(a) Geometric/exponential model: $P_{\mathbf{t+1}} = \lambda^t P_{\mathbf{t}}$ (Linear)

(b) Logistic model: $\boxed{\frac{\Delta P}{P} = r\,(1 - \frac{P}{K})}$ (nonlinear);

It can also be written in the form below as $\Delta P = P_{t+1} - P_t$

$$P_{t+1} = P_t\,(1 + r(1 - P_t/K))$$

- Nonlinear models are studied further graphically (cobweb method) or numerically (iterations).

- Linear stability analysis around the fixed/equilibrium point is done to determine the stability of every fixed point in the model.

Summary

- Fixed/equilibrium point is determined as follows:

  (a) If the model is written $P_{t+1} = F(P_t)$, *then* $P^* = F(P^*)$.

  (b) If the model is written as $\Delta P = G(P_t)$, *then* $G(P^*) = 0$.

  Note: P^* is the equilibrium point.

- To determine local stability of fixed/equilibrium points, two methods are used

(a) In Linear Perturbation method, $P_{t+1} = P^* + p_{t+1}$ *and getting*

$$p_{t+1} = r\, p_t\,.$$

Note: All the nonlinear terms are neglected

If $r < 1$ *(Perturbation dies, so stable) and*

if $r > 1$ *(Perturbation grows, so unstable)*

(b) First derivative method;

$|F'(P^*)| > 1$ (unstable) and

$|F'(P^*)| < 1$ (stable)

Problem-0

- Find the solution of the first order difference equation

(a) $x_{n+1} = \frac{1}{2} x_n$, $x_0 = 2$.

(b) $x_{n+1} = (\frac{1}{2} x_n) + 1$, $x_0 = 2$. (Home work)

Problem-1

1.1.7. Complete the following:

a. The models $P_t = k P_{t-1}$ and $\Delta P = r P$ represent *growing* populations when k is any number in the range ___ and when r is any number in the range ___.

H.W

b. The models $P_t = k P_{t-1}$ and $\Delta P = r P$ represent *declining* populations when k is any number in the range ___ and when r is any number in the range ___.

c. The models $P_t = k P_{t-1}$ and $\Delta P = r P$ represent *stable* populations when k is any number in the range ___ and when r is any number in the range ___.

Problem-2

1.1.10. A model is said to have a *steady state* or *equilibrium point* at P^* if whenever $P_t = P^*$, then $P_{t+1} = P^*$ as well.

a. Rephrase this definition as: A model is said to have a *steady state* at P^* if whenever $P = P^*$, then $\Delta P = \ldots$.

b. Rephrase this definition in more intuitive terms: A model is said to have a *steady state* at P^* if ....

c. Can a model described by $P_{t+1} = (1 + r)P_t$ have a steady state? Explain.

Problem-3

Table 1.3. *U.S. Population Estimates*

Year	Population (in 1,000s)
1920	106,630
1925	115,829
1930	122,988
1935	127,252
1940	131,684
1945	131,976
1950	151,345
1955	164,301
1960	179,990

1.1.14.

b. Using the data only from years 1920 and 1925 to estimate a growth rate for a geometric model, see how well the model's results agree with the data from subsequent years.

c. Rather than just using 1920 and 1925 data to estimate a growth parameter for the U.S. population, find a way of using all the data to get what (presumably) should be a better geometric model. (Be creative. There are several reasonable approaches.) Does your new model fit the data better than the model from part (b)?

No simple exponential model can fit this data

We can take the mean of all ratios to get an approximate estimate

Problem-4

1.2.5. Four of the many common ways of writing the discrete logistic growth equation are:

$$\Delta P = rP(1 - P/K), \qquad \Delta P = sP(K - P),$$

$$\Delta P = tP - uP^2, \qquad P_{t+1} = vP_t - wP_t^2.$$

Write each of the following in all four of these forms.

a. $P_{t+1} = P_t + .2P_t(10 - P_t)$

H.W

b. $P_{t+1} = 2.5P_t - .2P_t^2$

Problem-5 (Estimation of parameters)

Table 1.6. *Insect Population Values*

t	0	1	2	3	4	5	6	7	8	9	10
P_t	.97	1.52	2.31	3.36	4.63	5.94	7.04	7.76	8.13	8.3	8.36

1.2.7. If the data in Table 1.6 on population size were collected in a laboratory experiment using insects, would it be at least roughly consistent with a logistic model? Explain. If it is consistent with a logistic model, can you estimate r and K in $\Delta N = rN(1 - N/K)$?

Here K will be the carrying capacity where this population is converging to

Since P(t+1)/Pt = 1+r

Problem-6

1.2.8. Suppose a population is modeled by the equation

$$N_{t+1} = N_t + .2N_t(1 - N_t/200000)$$

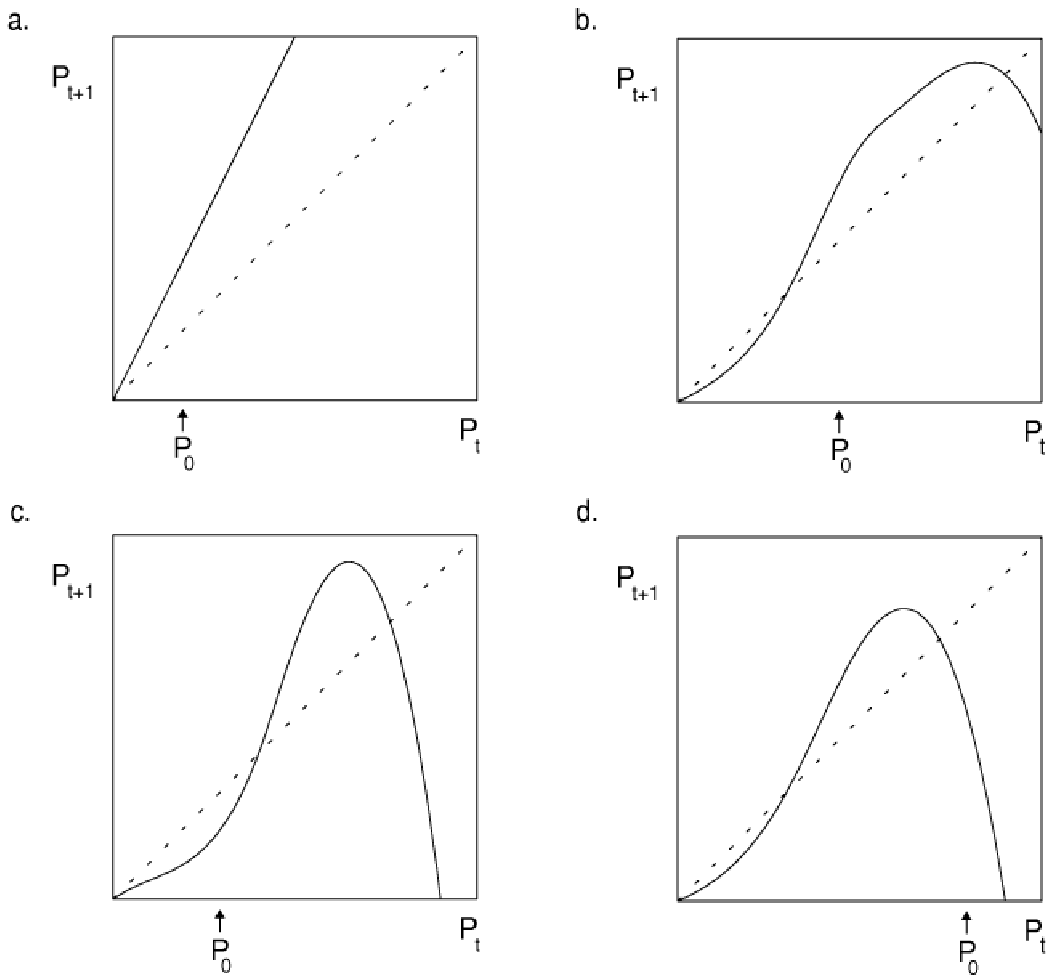
when N_t is measured in *individuals*.

a. Find an equation of the same form, describing the same model, but with the population measured in *thousands of individuals*. (*Hint*: Let $N_t = 1000M_t$, $N_{t+1} = 1000M_{t+1}$, and find a formula for M_{t+1} in terms of M_t.)

H.W

b. Find an equation of the same form, describing the same model, but with the population measured in units chosen so that the carrying capacity is 1 in those units. (To get started, determine the carrying capacity in the original form of the model.)

Problem-7 (already done, but clarification required)



Note:

Some cobweb plots are wrong in the solution manual.

Problems-8, 9, 10 & 11

1.3.6. For each of the following, determine the equilibrium points.

a. $P_{t+1} = 1.3P_t - .02P_t^2$

b. $P_{t+1} = 3.2P_t - .05P_t^2$

c. $\Delta P = .2P(1 - P/20)$

d. $\Delta P = aP - bP^2$

e. $P_{t+1} = cP_t - dP_t^2$.

1.3.7. For (a–e) of the preceding problem, algebraically linearize the model first about the steady state 0 and then about the other steady state to determine their stability.

H.W

1.3.8. Compute the equilibrium points of the model $P_{t+1} = P_t + rP_t(1 - P_t)$. Then use only algebra to linearize at each of these points to determine when they are stable or unstable.

H.W

1.3.9. (Calculus) Redo the preceding problem, but use derivatives to determine the stability of the equilibria of $P_{t+1} = P_t + rP_t(1 - P_t)$. You should, of course, get the same answers.