Data Analitiği

“Matematiksel Modelleme ve Tahmin”

Dr. Cahit Karakuş

Karar Destek Modellemesi

- Tahmine dayalı modeller, yüksek performanslı sistemlerde neredeyse her zaman karar vermenin temelini oluşturur.

Matematiksel modelleme nedir?

- Gerçek bir durumu kabul edilebilir bir doğruluk düzeyinde tanımlayan bir dizi matematiksel denklemin geliştirilmesi ve çözümü.Gerçek bir durumda ne olacağını tahmin etmek için kullanılırNot: Burada tamamen nitel/tanımlayıcı modellerden ziyade nicel modellerle ilgileniyoruz.

Example – A simple population model (the "logistic map")

Let N_i be the population in year i
A crude model for population growth is that
$N_i = aN_{i-1}$
When $a>1$, the population increases (unboundedly) each year. When $a<1$, the population decreases each year until it reaches zero (extinction).
This is obviously an oversimplification. We need to account for overcrowding and limited resources i.e. we expect the value of parameter a to vary with the population N.
Now have

$$N_i = a(N_{i-1})N_{i-1}$$

Lets assume that $a(N) = \alpha(1-N)$ for N in the interval [0,1]

Example – A simple population model (the "logistic map")

$N_i = \alpha(1-N_{i-1})N_{i-1}$ "Logistic Map"

Note that this is a very much simplified model. Are these simplifications justified ? That is, how "good" is this model.

-depends on what we want to use it for.

-can compare predictions of model with observed behaviour to validate model/establish accuracy.

Availability of observations places fundamental limit on quality of model.

What can we use the model for ?

The structure of the model reflects our understanding of the structure of the system

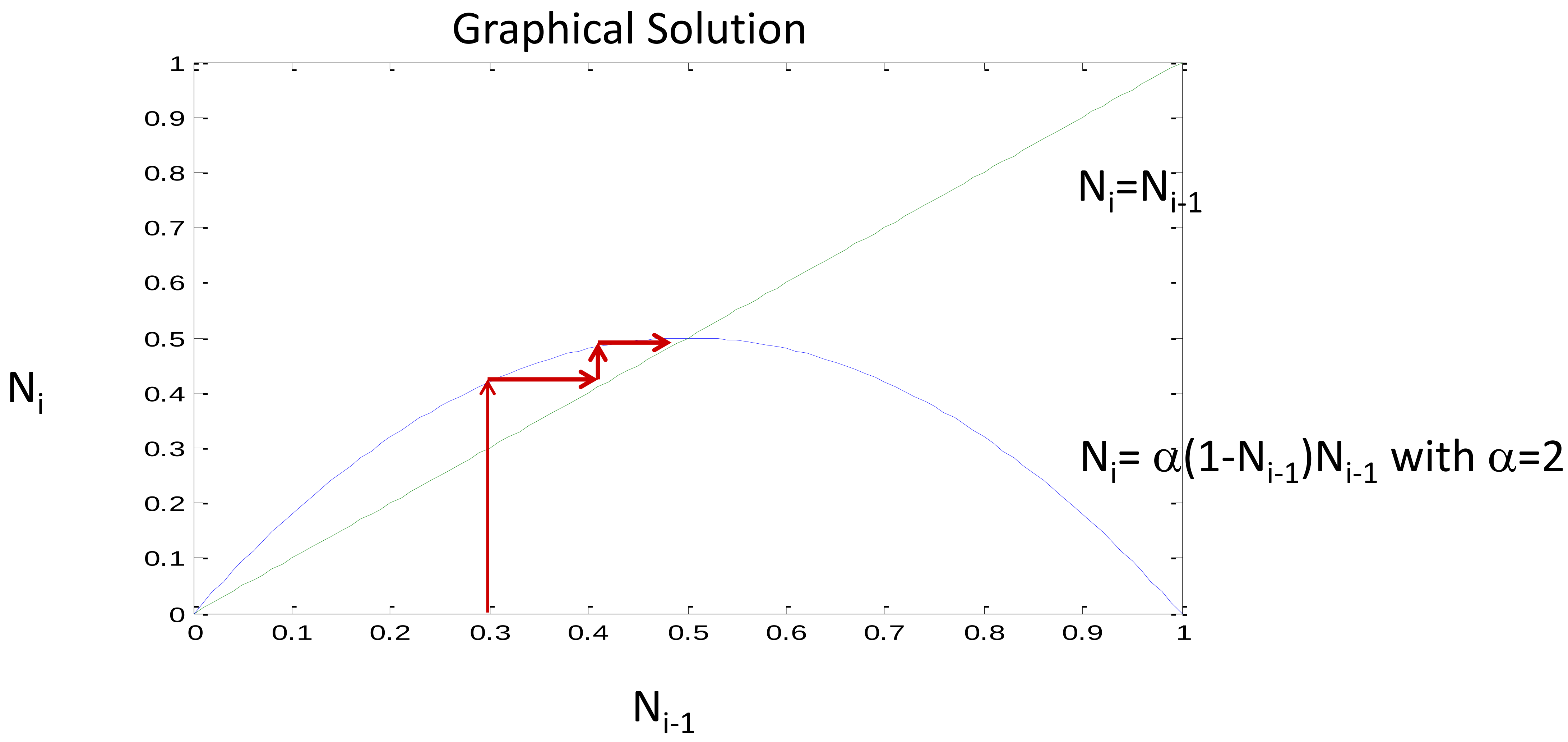
-a model organises our experiences and observations

We can solve the model equations to make predictions

- decision support

Example – A simple population model (the "logistic map")

We can solve the model equations to make predictions …

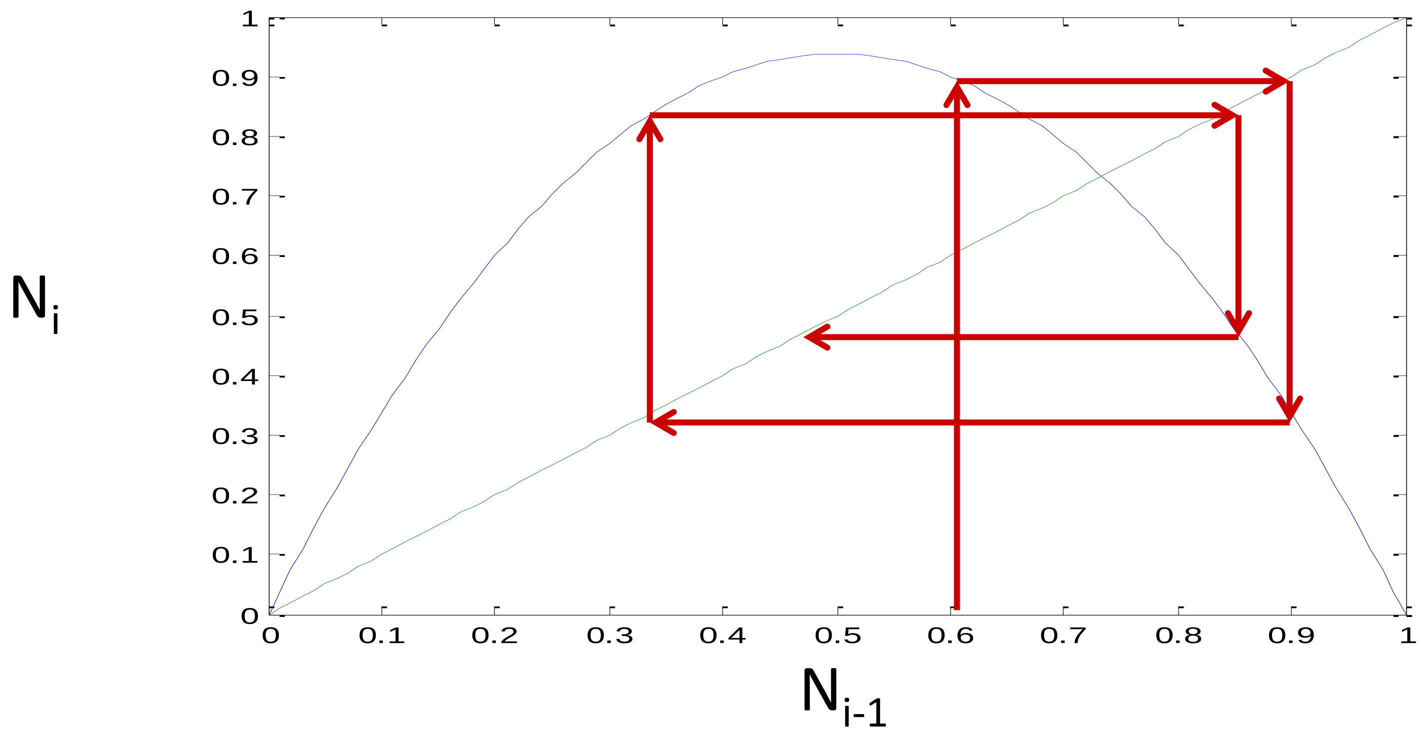


Example – A simple population model (the "logistic map")

With α=2, a stationary point exists to which all solutions in interval [0,1] are eventually attracted.

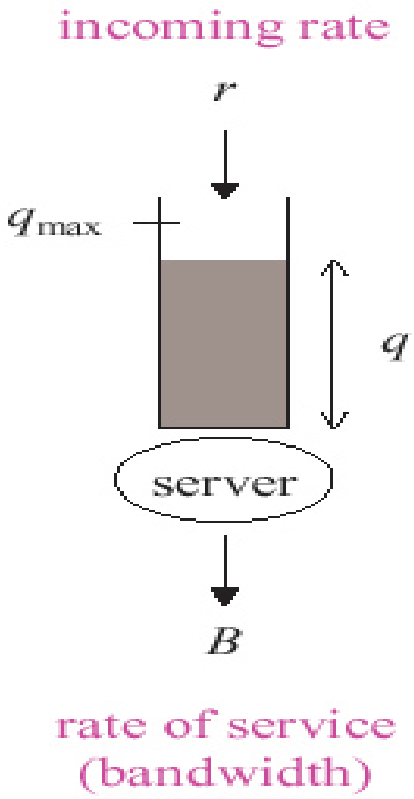
NB: "stationary point" = "equilibrium point" = "steady-state solution"

This is not always the case. For example, consider when α=3.75



The population never settles down to a constant value

Example: Server with buffer/queue



Queue length, q. Service rate, B packets/s.

Packets arrive at times $t_0, t_1, t_2, \ldots$

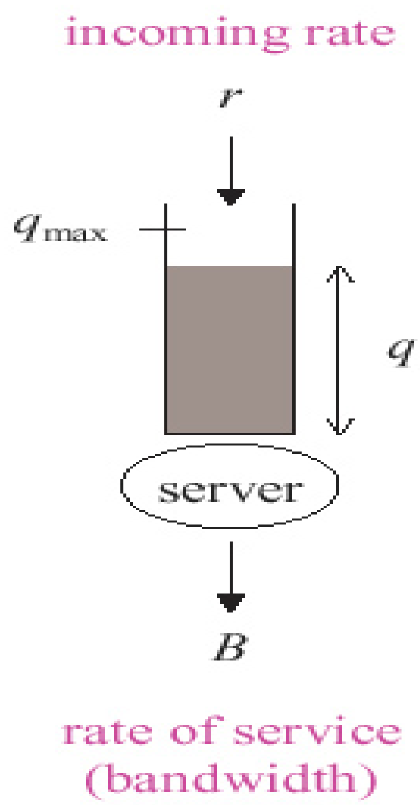
Ignoring servicing of packets just now, when a packet arrives we have

$$q(t_n)=\min(q(t_{n-1})+1,\ q_{max})$$

Now, interval between two packets is $t_n - t_{n-1}$. During this interval $\lfloor B(t_n-t_{n-1}) \rfloor$ packets are serviced i.e. removed from the queue. The queue size cannot fall below zero. So, we have as our model

$$Q=\min(q(t_{n-1})+1-\lfloor B(t_n-t_{n-1}) \rfloor,\ q_{max})$$
$$q(t_n)=\max(Q,0)$$

Example: Server with buffer/queue and acknowledgement

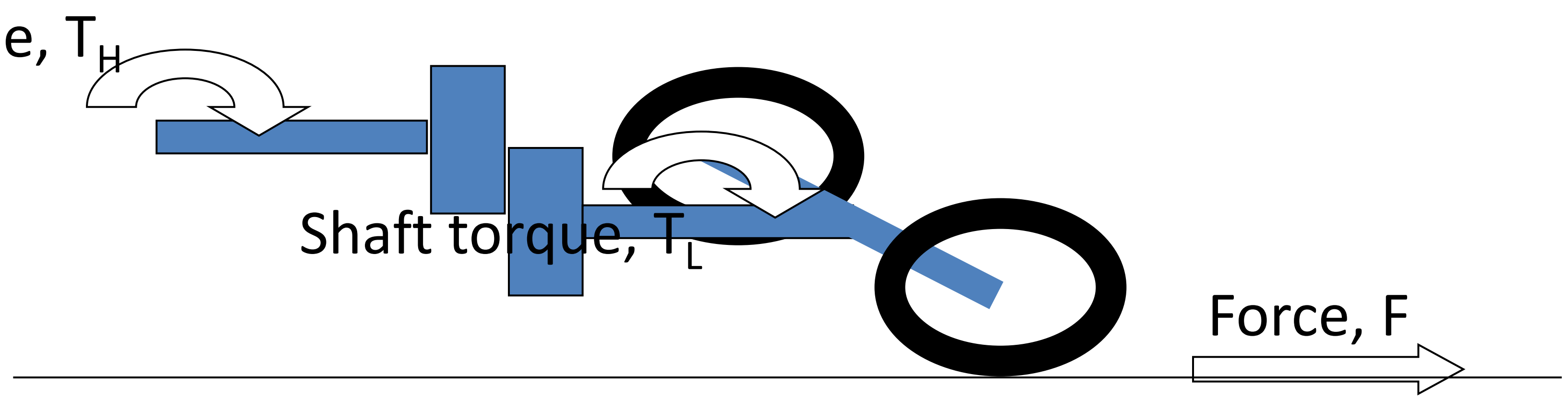


Now let's model the behaviour of the source. Suppose we have one source and that it sends a new packet in response to the server signalling that it has finished servicing a packet. Also suppose that its time T for the packet to travel from the source to the queue.

1. Send first packet at time $t_{0.}$
2. Packet arrives at queue at time t_o+T
3. Service rate is B packets/s, so at time t_o+T+1/B server signals that packet has been serviced.
4. Send second packet – time is now $t_1=t_0+T+1/B$
5. Packet arrives at queue at time t_o+2T+1/B

Queue now doesn't overflow, but server is idle for time T between packets arriving. Can we do better ?

Example: Vehicle transmission



$T_L = NT_H$, N is gearbox ratio

Force exerted by wheel is $\mu T_L R$, with μ friction coeff, R radius of wheel

Air creates drag force $-\lambda v$, with v the velocity of vehicle

Newton's Law: Force = mass*acceleration

$F = \mu N T_H R - \lambda v = m\, a$

Noting that a=dv/dt, we have the following model for the speed of the vehicle.

$m\, dv/dt = \mu N T_H R - \lambda v$

Example: Vehicle transmission

Suppose the vehicle has two gears, N_1 and N_2. The gear used is selected by an automatic transmission. The model is then

$$m\, dv/dt = \begin{cases} \mu N_1 T_H R - \lambda v & \text{in gear 1} \\ \mu N_2 T_H R - \lambda v & \text{in gear 2} \end{cases}$$

+ a model of the the decision making process used by the automatic transmission.

Course Outline

How do we derive a model for a system ?
How do we extract information from it (esp. how do we obtain quantitative solutions and analyse their properties) ?

- this course is structured around these questions.

- Introduce taxonomy of models. It turns out that most systems can be modelled using a fairly small set of model structures.

- Study solutions, esp. numerical solutions/simulations

- Can derive models from first principles or learn model from observations (or more usually by a combination of both approaches). We will not cover first principles modelling as very application specific. But will introduce machine learning approaches (including probabilistic reasoning ideas).

A taxonomy of mathematical models

- Difference Equations
- Differential Equations
- Hybrid

Can combine these two classifications e.g. linear differential equations

- Linear
- Nonlinear

- Time-invariant
- Time-varying

Other aspects of models can also be usefully classified, but not pursued here.
Especially deterministic/stochastic models – stochastic models not covered in this course.

A taxonomy of mathematical models

- Difference Equations
- Differential Equations
- Hybrid

Simple Example of a Difference Equation:

$y(k) = a\, y(k-1),\ k=1,2,\ldots$

This is equivalent to the (infinite) set of equations:

$y(1)=ay(0)$
$y(2)=ay(1)$
$y(3)=ay(2)$
etc.

If have observations of $y(0)$, $y(1)$, etc, then this defines a relation between these observations. If $y(1)$, $y(2)$ etc are unknown, the equations can be solved to find them.

Difference Equations

Logistic Map is another example of a difference equation

$y(k)=\alpha(1-y(k-1))y(k-1),\ \ k=1,2,\ldots$

Can also include an external input u in the difference equation, e.g.

$y(k) = a\, y(k-1)+bu(k-1),\ k=1,2,\ldots$

This is equivalent to the (infinite) set of equations:

$y(1)=ay(0)+bu(0)$
$y(2)=ay(1)+bu(1)$
$y(3)=ay(2)+bu(2)$
etc.

Difference Equations

Definition:

Suppose there is a defined sequence of values y(k), k=0,1,2,… (e.g. representing values observed at equally-spaced time points). A difference equation is an equation relating the value y(k) to other values y(i), $i \neq k$.

A difference equation is said to be causal when y(k) is related to values y(i) with $i<k$.

$$y(k)=f(y(k-1),y(k-2),\ldots,y(k-n),\underbrace{u(k-1),u(k-2),\ldots,u(k-m)}_{u \text{ is an external input}})$$

where m,n are some constants.

NB: We write y(k)=f(y(k-1),y(k-2),…) but could equally well write this as y(k+1)=f(y(k),y(k-1),…), and this is often done.

Difference Equations

A difference equation is said to be linear when the function f on the RHS is a linear function

i.e $y(k)=a_1(k)y(k-1) + a_2(k)\, y(k-2)+\ldots+a_n(k)\, y(k-n)$
$+b_1(k)\, u(k-1)+b_2(k)\, u(k-2)+\ldots+b_m(k)\, u(k-m))$

where $a_1(k)$, $a_2(k)$, …,$a_n(k)$ and $b_1(k)$,$b_2(k)$, …, $b_m(k)$ are time-varying parameters. When these parameters are constants (do not vary with time), the model is said to be linear time-invariant.

- Difference Equations
- Differential Equations
- Hybrid

- Linear
- Nonlinear

- Time-invariant
- Time-varying

+ → Linear Time-Invariant Difference Equations

Difference Equations

A solution to a difference equation is a function $y(k)$ that satisfies the equation. Solutions are readily derived by recursion.

e.g. for $y(k)=ay(k-1)$ we have that

$y(1)=ay(0)$

$y(2)=ay(1)=a^2y(0)$

$y(3)=ay(2)=a^3y(0)$

etc.

i.e. a solution is $y(k)=a^k y(0)$

NB: We need to specify $y(0)$ in order to solve this equation. This is called the initial condition for the equation. We need to specify both the equation and its initial condition in order to define a solution.

Difference Equations – Initial Conditions

More generally,

$$y(k)=f(y(k-1),y(k-2),\ldots,y(k-n),\ u(k-1),u(k-2),\ldots,u(k-m))$$

We assume that the input values $u(k)$ are defined beforehand. We must also specify $y(0)$, $y(1)$, …, $y(n-1)$ in order to define a solution – in general the initial condition must specify n values.

Difference Equations – Existence & Uniqueness of Solutions

- Note that a difference equation need not have any solution.

e.g. $y(k)^2=-(1+y(k-1)^2)$

has no solution since $y(k)^2$ can never be negative.

- Also, even when a solution exists, it need not be unique i.e. there may exist many solutions.

e.g. $\sin y(k) = y(k-1)$

Generally, however, a model of a physical system can be expected to possess a solution which is unique.

Difference Equations – Solutions to Linear Time-Invariant Models

Recall that linear time-invariant difference equations have the form

$y(k)=a_1y(k-1) + a_2y(k-2)+\ldots+a_ny(k-n)$

$\qquad +b_1u(k-1)+b_2u(k-2)+\ldots+b_mu(k-m))$

Consider the simplest system $y(k)=ay(k-1)$ "first-order system"

We have that $y(1)=ay(0)$

$y(2)=ay(1)=a^2y(0)$

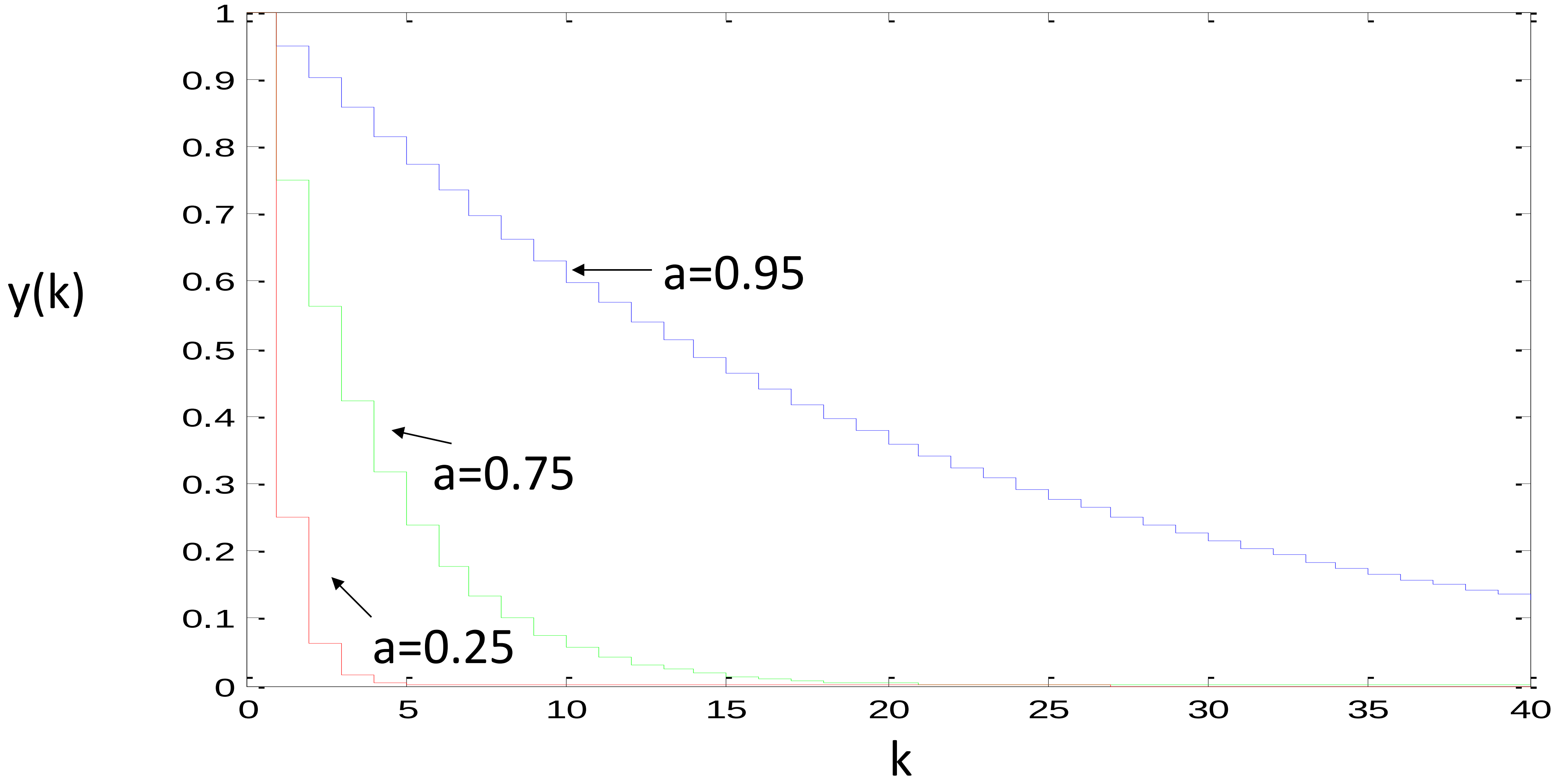
$y(3)=ay(2)=a^3y(0)$

etc

So the solution is $y(k)=a^ky(0)$. Note that the solution behaves as an exponential – we can rewrite it as $y(k)=\exp(k \log a)y(0)$

- For $a<1$, $y(k)\rightarrow 0$ as $k\rightarrow\infty$ - system is said to be stable
- For $a>1$, $y(k)\rightarrow\infty$ as $k\rightarrow\infty$ - system is said to be unstable
- For $a=1$, solution neither grows of decays
  - system is said to be critically stable

Also, rate of convergence/divergence varies with the value of *a*



Consider now $y(k)=a_1y(k-1) + a_2y(k-2)$ "second-order system"

By analogy to the first-order case, try a solution of the form $y(k)=\lambda^k y(0)$ where λ is some (as yet unknown) constant. Then we need,

$$\lambda^k y(0)=a_1 \lambda^{k-1}y(0)+a_2 \lambda^{k-2}y(0)$$

i.e. $\lambda^k - a_1 \lambda^{k-1} - a_2 \lambda^{k-2} = 0$

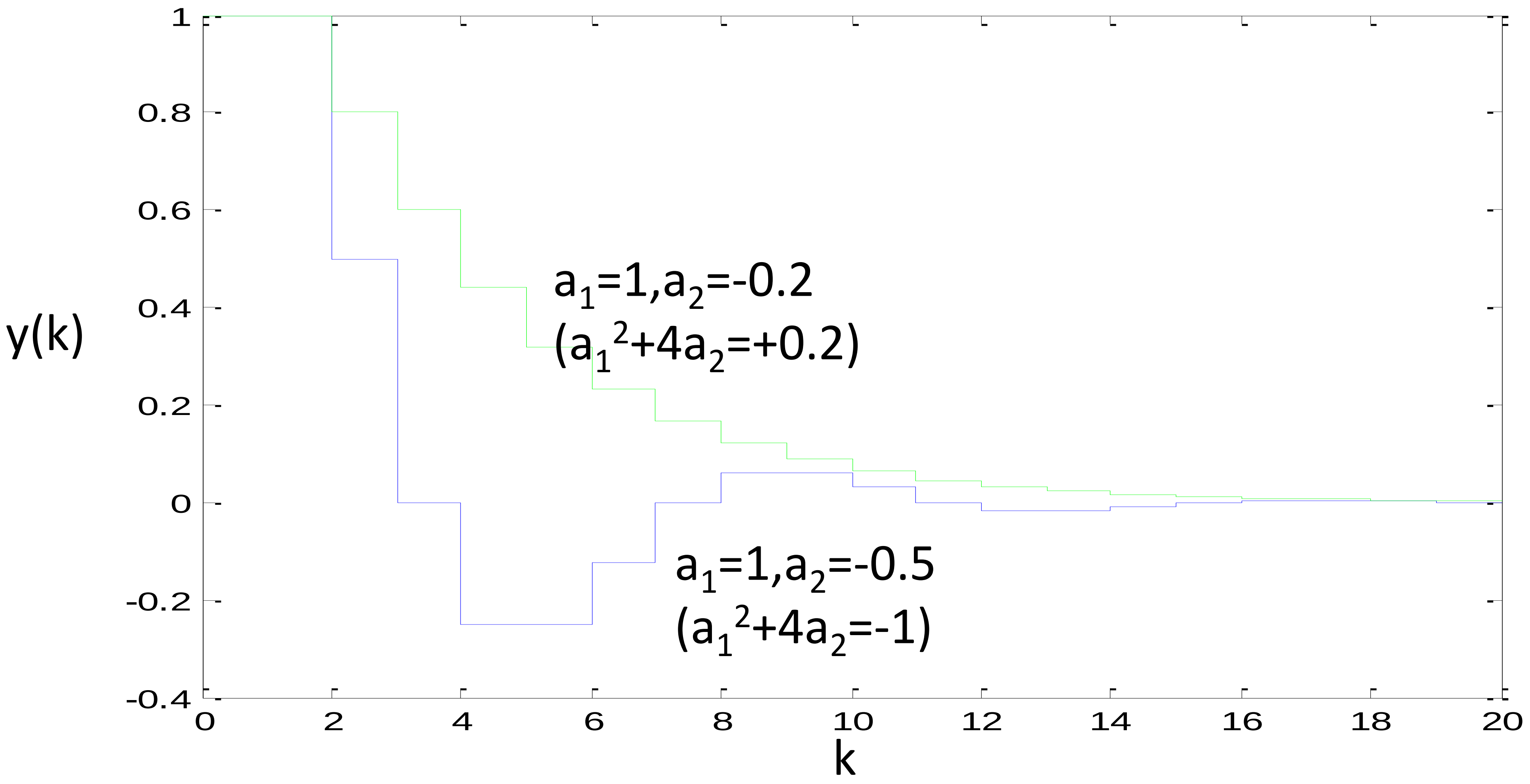
Dividing through by λ^{k-2} gives

$$\lambda^2 - a_1 \lambda - a_2 \lambda = 0$$

i.e. $\lambda= a_1/2 \pm\sqrt{(a_1^2+4a_2)}/2$

We can work this through to derive an explicit solution. We won't do this though. Observe that the situation where $a_1^2+4a_2 < 0$ (and so λ is complex valued) looks like its going to be different from when $a_1^2+4a_2 < 0$ (and so λ is real valued).

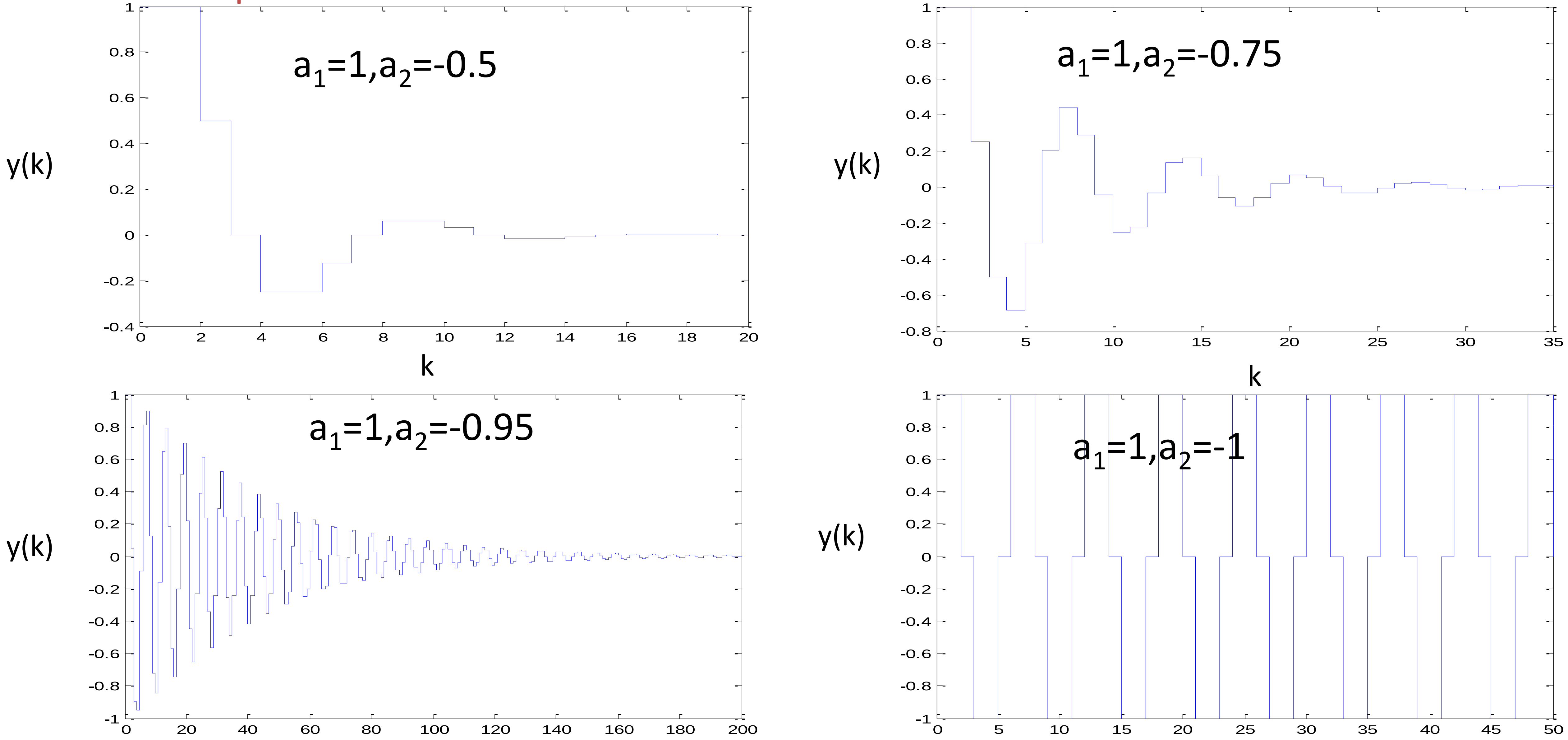
Consider now $y(k)=a_1y(k-1) + a_2y(k-2)$ "second-order system"



$a_1^2+4a_2 < 0$
"underdamped"

$a_1^2+4a_2 > 0$
"overdamped"

Difference Equations – Solutions to Linear Time-Invariant Models



NOTE:

$\lambda = a_1/2 \pm \sqrt{(a_1^2+4a_2)}/2$

Similarly to first-order case,

- For $|\lambda|<1$, $y(k) \to 0$ as $k \to \infty$ - system is stable
- For $|\lambda|>1$, $y(k) \to \infty$ as $k \to \infty$ - system is unstable
- For $|\lambda|=1$, solution neither grows of decays
  – system is critically stable

When λ is real valued, system is overdamped (with special case called critically damped when $a_1^2+4a_2=0$)

-solution to system is the sum of pure exponentials

When λ is complex valued, system is underdamped

-solution to system is oscillatory, with envelope that decays exponentially for stable systems, grows exponentially for unstable systems.

Difference Equations – Equilibria of Linear Time-Invariant Models

For stable systems, the solution converges to a final value as $k \to \infty$. This is called the equilibrium point of the system (also called stationary point or steady-state value).

Linear time-invariant difference equation:

$$y(k)=a_1 y(k-1) + a_2 y(k-2)+\ldots+a_n y(k-n)$$
$$+b_1 u(k-1)+b_2 u(k-2)+\ldots+b_m u(k-m))$$

When the input u is zero, at an equilibrium point y_∞ we must have: $y_\infty=a_1\, y_\infty+a_2\, y_\infty+\ldots+a_n\, y_\infty$
i.e. $y_\infty=0$

When input is non-zero, the equilibrium point will depend on the input. E.g. say u(k)=u, a constant value. Then

$$y_\infty=a_1\, y_\infty+a_2\, y_\infty+\ldots+a_n\, y_\infty+b_1 u+b_2 u+\ldots+b_m u$$

i.e. $y_\infty= u\, (b_1+b_2+\ldots+b_m)/(a_1+a_2+\ldots+a_n)$

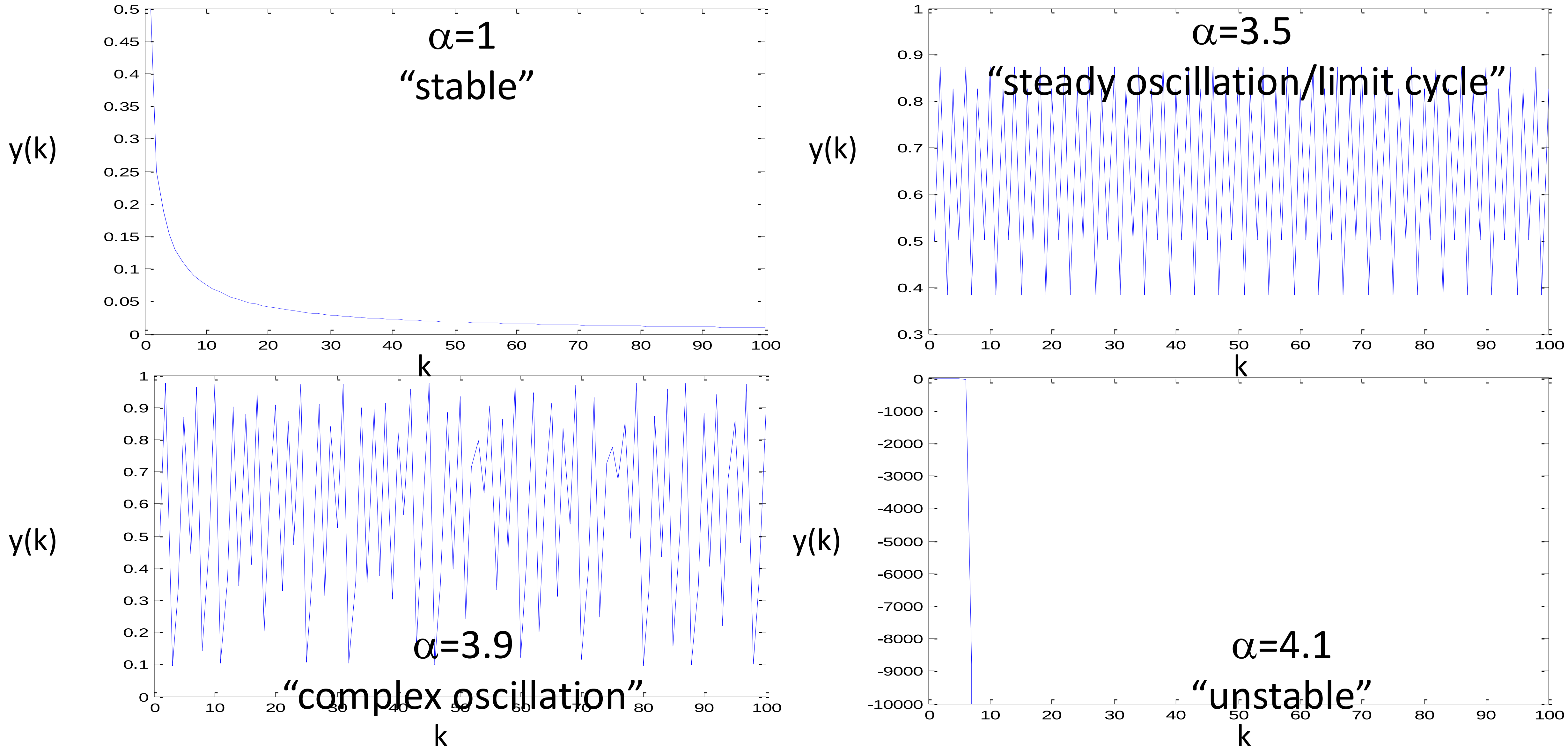
Difference Equations – Solutions to Nonlinear Models

For linear difference equations, in qualitative terms only a small number of types of solution can exist (stable, unstable, overdamped, underdamped etc).

For nonlinear difference equations, the situation is much richer.

e.g. depending on the value of the parameter α, the Logistic Map $y(k)= \alpha(1-y(k-1))y(k-1)$ not only has solutions which are stable and unstable but also has steady oscillatory solutions and chaotic solutions (complex oscillations) - these are both types of equilibrium solution which do not have just a single value y_∞.

Difference Equations – $y(k)= \alpha(1-y(k-1))y(k-1)$



Differential Equations

A simple example is

$dy(t)/dt = ay(t)$

Verify that the solution to this differential equation is:

$y(t)=\exp(at)\, y(0)$

-compare with the first-order difference equation $y(k)=ay(k-1)$
which has solution $y(k)=a^k y(0)=\exp(k \log(a))\, y(0)$.

Another example:

$$\frac{d^2y}{dt^2} + \sin y \frac{dy}{dt} = \cos t$$

Can also include an external input u in the differential equation, e.g.

dy(t)/dt = a y(t)+bu(t)

Definition:

Suppose there is a function y(t) defined on an interval $[t_o,t_1]$. A differential equation is an equation relating the value y(t) to some of its derivatives.

In general, a differential equation is of the form:

Differential Equations

A differential equation is said to be linear when the function f on the RHS is a linear function

i.e
$$\frac{d^n y(t)}{dt^n} = a_1(t)y(t) + a_2(t)\frac{dy(t)}{dt} + a_3(t)\frac{d^2 y(t)}{dt^2} + \ldots + a_n(t)\frac{d^{n-1} y(t)}{dt^{n-1}}$$
$$+ b_1(t)u(t) + b_2(t)\frac{du(t)}{dt} + \ldots + b_{m+1}(t)\frac{d^m u(t)}{dt^m}$$

where $a_1(t)$, $a_2(t)$, …,a_n (t) and b_1 (t),b_2 (t), …, b_m (t) are time-varying parameters. When these parameters are constants (do not vary with time), the model is said to be linear time-invariant.

- Difference Equations
- Differential Equations
- Hybrid
- Linear
- Nonlinear
- Time-invariant
- Time-varying

\+ → Linear Time-Invariant Differential Equations

A solution to a differential equation is a function y(t) that satisfies the equation.

E.g. the solution to $dy(t)/dt = ay(t)$ is $y(t)=\exp(at)\, y(0)$

Similarly to difference equations, we need to specify the initial condition y(0) in order to solve this equation i.e. we need to specify both the equation and its initial condition in order to define a solution.

E.g. the differential equation:

$$\frac{d^2y}{dt^2} = 0$$

has general solution of the form $y(t)=A+Bt$ where A,B are some constants. To find the values of these constants we use

$$y(0)=A, \quad dy(0)/dt=B$$

Also, similarly to the difference equation $y(k)=ay(k-1)$, the solution to $dy(t)/dt = ay(t)$ is $y(t)=\exp(at)\, y(0)$

i.e.

- For $a<0$, $y(t)\rightarrow 0$ as $t\rightarrow\infty$ - system is said to be stable
- For $a>0$, $y(t)\rightarrow\infty$ as $t\rightarrow\infty$ - system is said to be unstable
- For $a=0$, $y(t)=y(0)$ - system is said to be critically stable

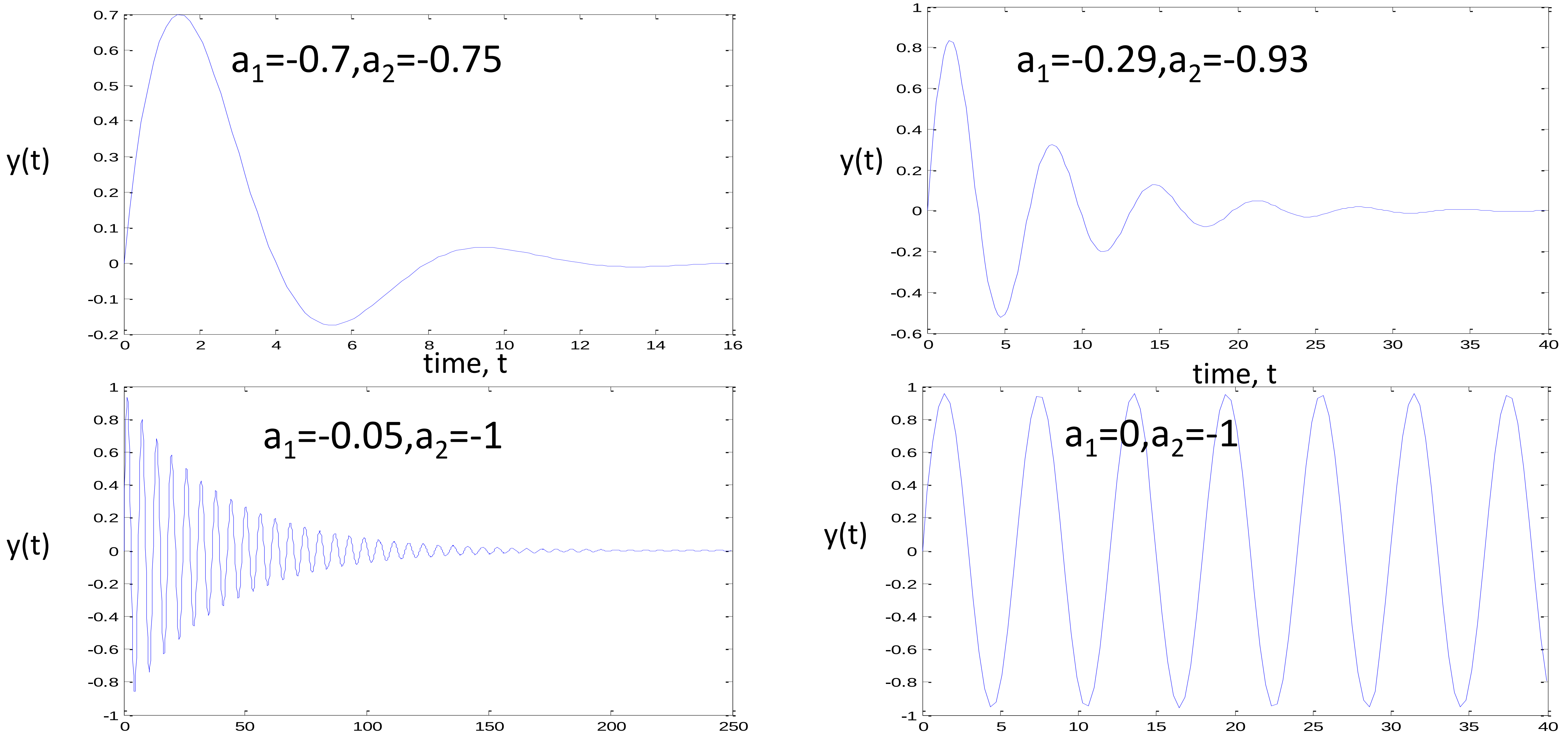
Similarly to the second-order linear difference equation, the second-order linear differential equation

$$\frac{d^2y}{dt^2} = a_1 y + a_2 \frac{dy}{dt}$$

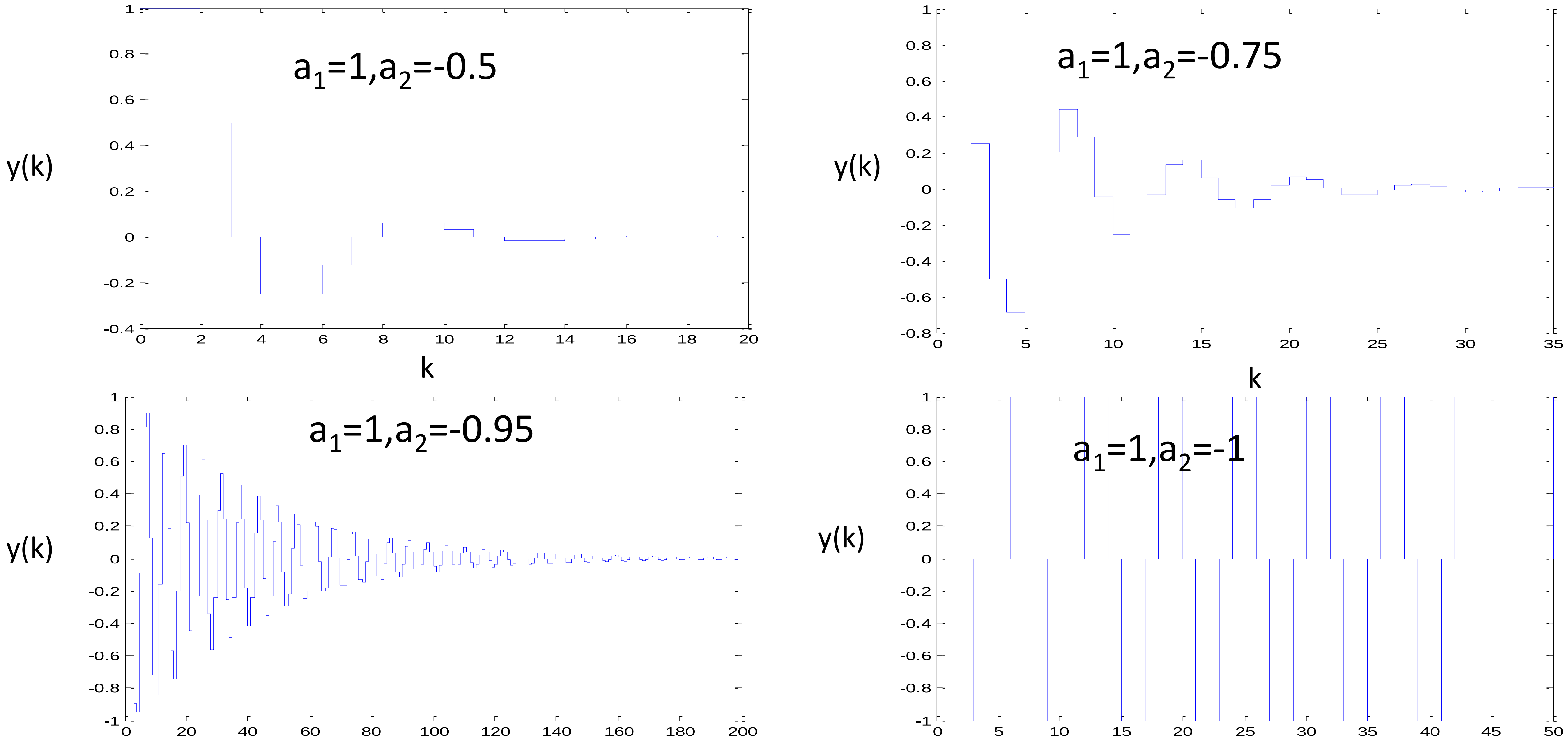
exhibits overdamped, underdamped and critically damped responses depending on the values of a_1 and a_2.

Second-order Linear Differential Equation

$$\frac{d^2y}{dt^2} = a_1 y + a_2 \frac{dy}{dt}$$



Compare with 2nd order difference eqn $y(k)=a_1y(k-1) + a_2y(k-2)$



Linear difference and differential equations are closely related.
NOTE: This is generally only true for linear equations. Nonlinear equations seem to be fundamentally different,
e.g. chaos can exist in first-order difference equations (such as the logistic map), but not in first-order differential equations (we need to go to at least third order to find a differential equation which exhibits chaos).

Recall definition of derivative:

$$\frac{dy(t)}{dt} = \lim_{h \to 0} \frac{y(t+h) - y(t)}{h}$$

This suggests that a derivative might be approximated by the finite difference:

$$\frac{dy(t)}{dt} \approx \frac{y(t+h) - y(t)}{h}$$

for small h. This is called the Euler approximation. We expect that as h is made smaller, the accuracy of the approximation improves.

Consider the first-order linear differential equation:

$$\frac{dy(t)}{dt} = a_1(t)y(t)$$

and now replace the exact derivative by its Euler approximation. We have

$$\frac{y(t+h) - y(t)}{h} = a_1(t)y(t)$$

i.e. $y(t+h)=y(t)+ha_1(t)y(t)$

Considering the time instants kh, k=0,1,2,… we have

$$y((k+1)h)=y(kh)+ha_1(kh)y(kh)$$

which we can write as

$$y(k+1)=y(k)+ha_1(k)y(k)$$

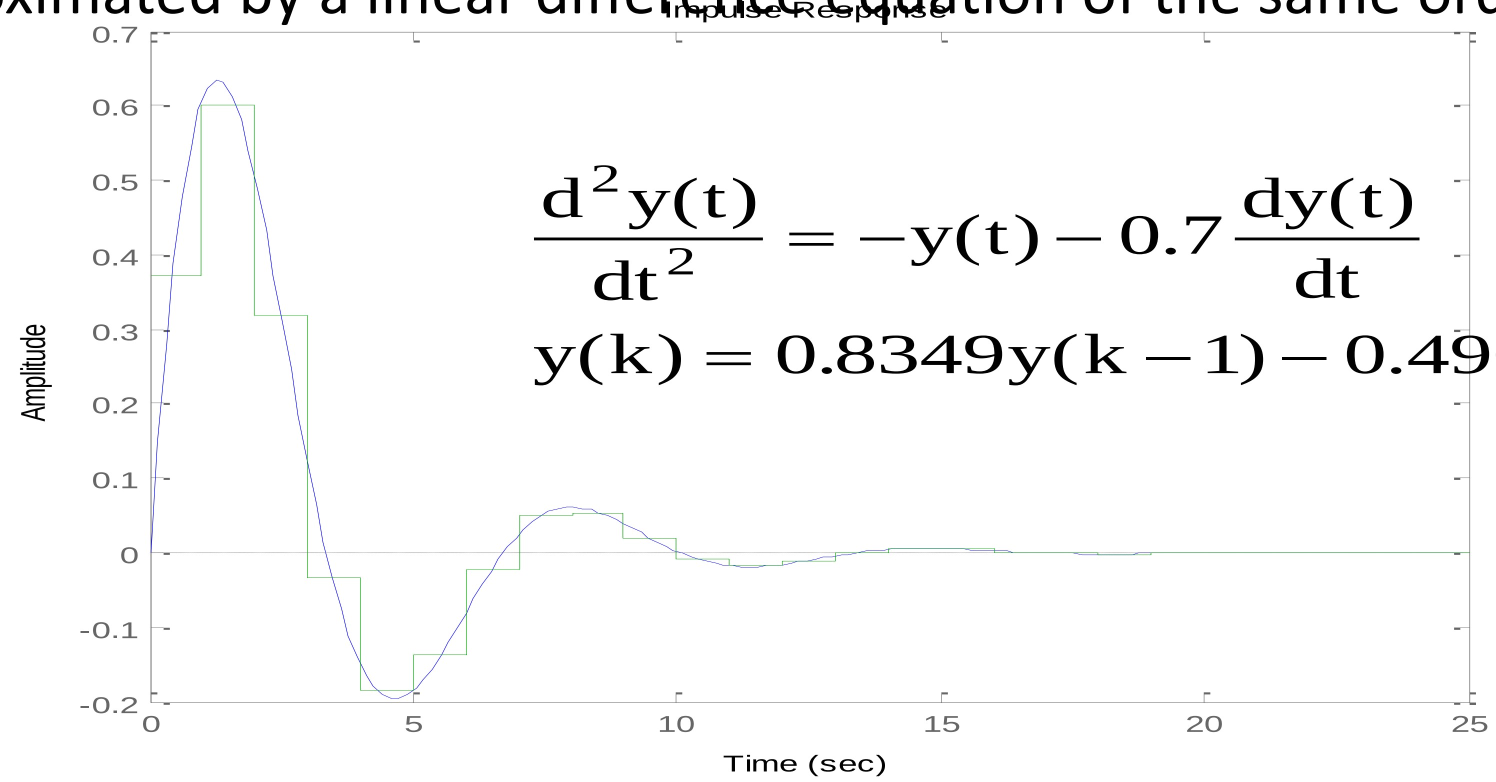
We have converted our first-order differential equation into a first-order difference equation. This conversion is only approximate, but the approximation becomes arbitrarily accurate as h is made small.

More generally, a linear differential equation:

$$\frac{d^n y(t)}{dt^n} = a_1(t)y(t) + a_2(t)\frac{dy(t)}{dt} + a_3(t)\frac{d^2y(t)}{dt^2} + \ldots + a_n(t)\frac{d^{n-1}y(t)}{dt^{n-1}}$$
$$+ b_1(t)u(t) + b_2(t)\frac{du(t)}{dt} + \ldots + b_{m+1}(t)\frac{d^m u(t)}{dt^m}$$

can be approximated by a linear difference equation of the same order.

Example



$$\frac{d^2y(t)}{dt^2} = -y(t) - 0.7\frac{dy(t)}{dt}$$

$$y(k) = 0.8349y(k-1) - 0.4966y(k-2)$$

The Concepts of a System

System: is any collection of interacting elements that operate to achieve some goal.