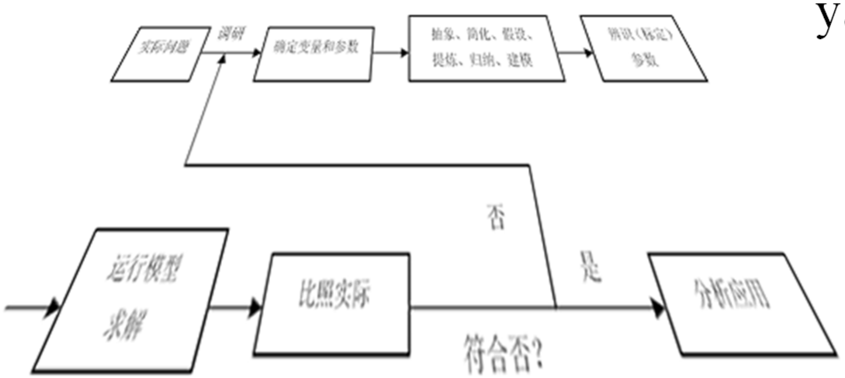
数据建模分析:
课程简介

姚 远

yaoyuan@shu.edu.cn



上海大学机自学院

2020/9/24

大纲

- 课程说明
- 考核指标
- 应用背景
- 基本概念
- 建模方法及原则

课程说明

- **课程内容**
  - 建模理论
  - 数据及处理方法
  - 实用技术及工具
- **考核指标**
  - 讲座7次(10%)
  - 1项作业(10%)
  - 作业课堂报告1次(10%)
  - 考试(70%)

https://rmec.gitee.io/courses/dam/2020Au/index.html

课程目标

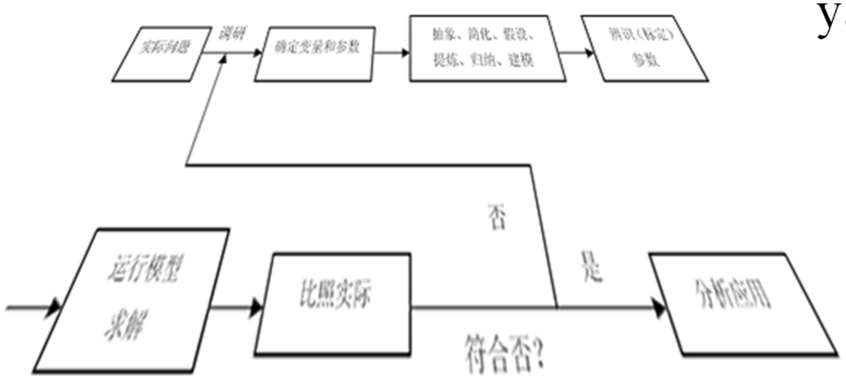
目的：数据的建模与分析过程贯穿于科研与工程的整个过程中

- 掌握系统建模与分析的基本概念
- 掌握数据分析的基本概念与处理方法
- 了解不同领域中模型表达和相关算法
- 了解现代人工智能建模的各类方法
- 掌握基本的程序设计能力

数据建模与分析:
建模仿真基础

姚 远

yaoyuan@shu.edu.cn



上海大学机自学院

2020/9/24

基本概念

基本概念

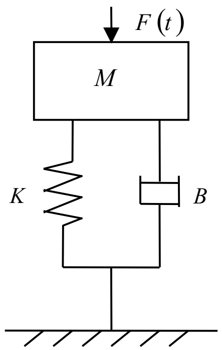
- 系统的定义——按照某些规律结合起来，互相作用、互相依存的所有实体的集合或总体。

- 描述系统的“三要素”：实体、属性、活动
- 实体——确定系统的构成和系统的边界。
- 属性——描述每一个实体的特征（即描述变量）。
- 活动——定义了系统内部实体之间的相互作用，从而确定了系统内部发生变化的过程。

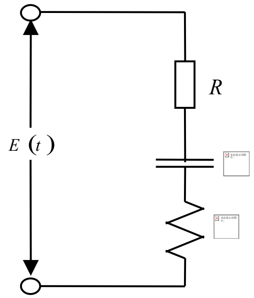
基本概念

系统模型的性质：

- **相似性**——相似的物理属性和数学描述。
- **简单性**——忽略次要因素和非可测变量的影响。
- **多面性**——基于研究目的不同反应系统的不同侧面。



$$M\frac{d^2x}{dt^2}+B\frac{dx}{dt}+Kx=F(t)$$



$$L\frac{d^2q}{dt^2}+R\frac{dq}{dt}+\frac{q}{C}=E(t)$$

基本概念

系统的分类：

(1) 工程领域: 机械、航空航天、电力、冶金、化工和电子等

非工程领域: 交通管理、生产调度、库存控制、生态环境和社会经济等

(2) 根据数学方法:

初等模型、微分方程模型、优化模型、线性或非线性模型

(3) 根据理解程度

白箱、黑箱、灰箱

基本概念

系统的分类：

(4) **CVDS** Continuous Variable Dynamic Systems)连续变量动态系统

DEDS (Discrete Event Dynamic Systems) 离散事件动态系统

HDS (Hybrid Dynamic Systems)

混合动态系统

根据系统性质、实际问题等等

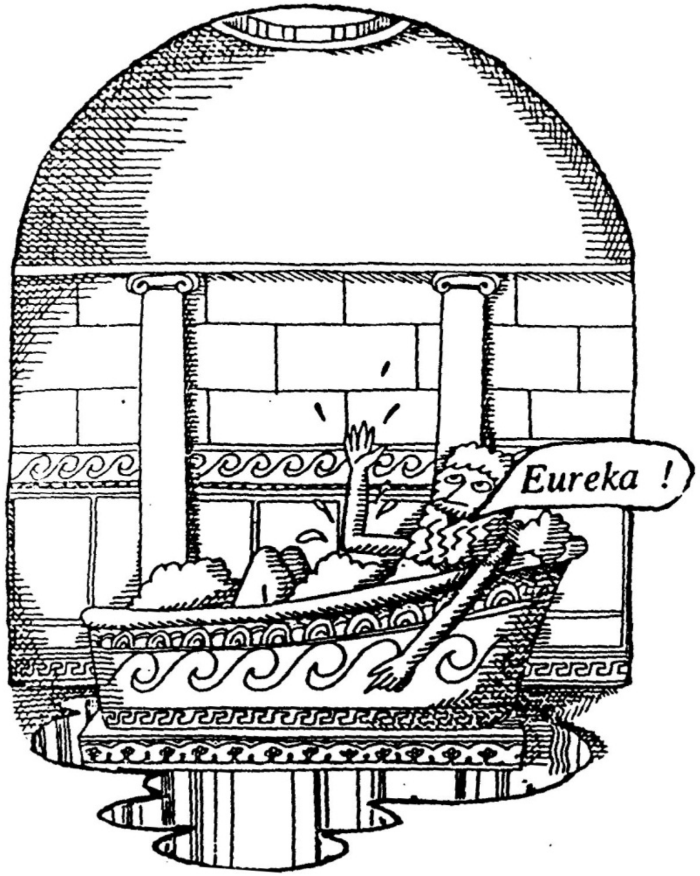
基本概念

连续系统
Continuous Variable
Dynamic Systems (CVDS)

从?至今

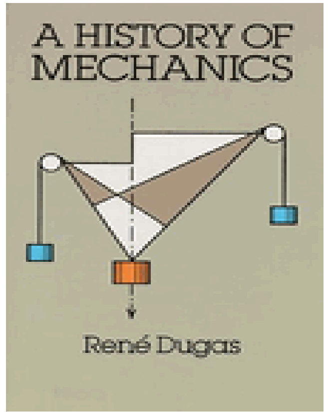
A History of Mechanics, R. Dugas, 1935

From



ancient Greeks

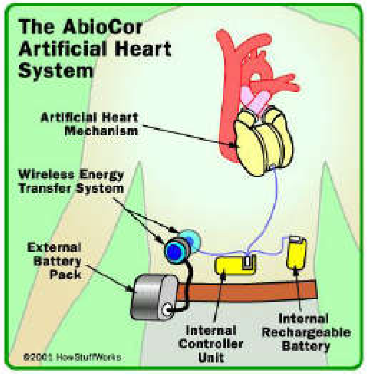
To

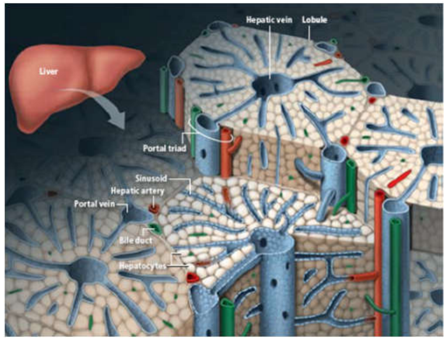


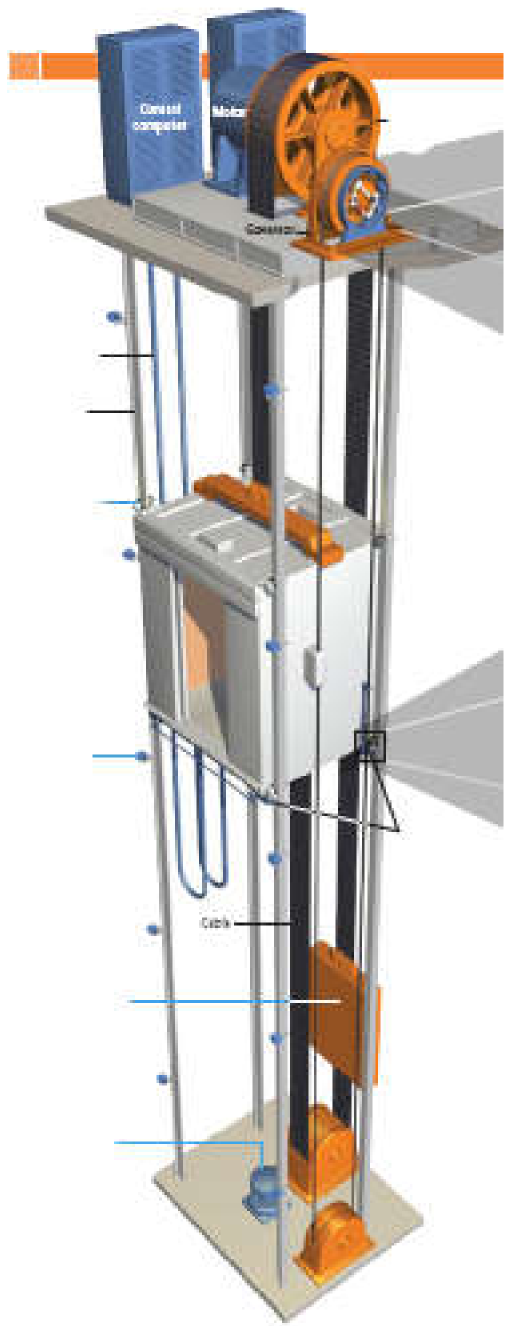
20 Century

$$i\frac{h}{2\pi}\frac{\partial\psi}{\partial t}=\mathcal{H}\psi$$

物理系统——时间连续、状态连续





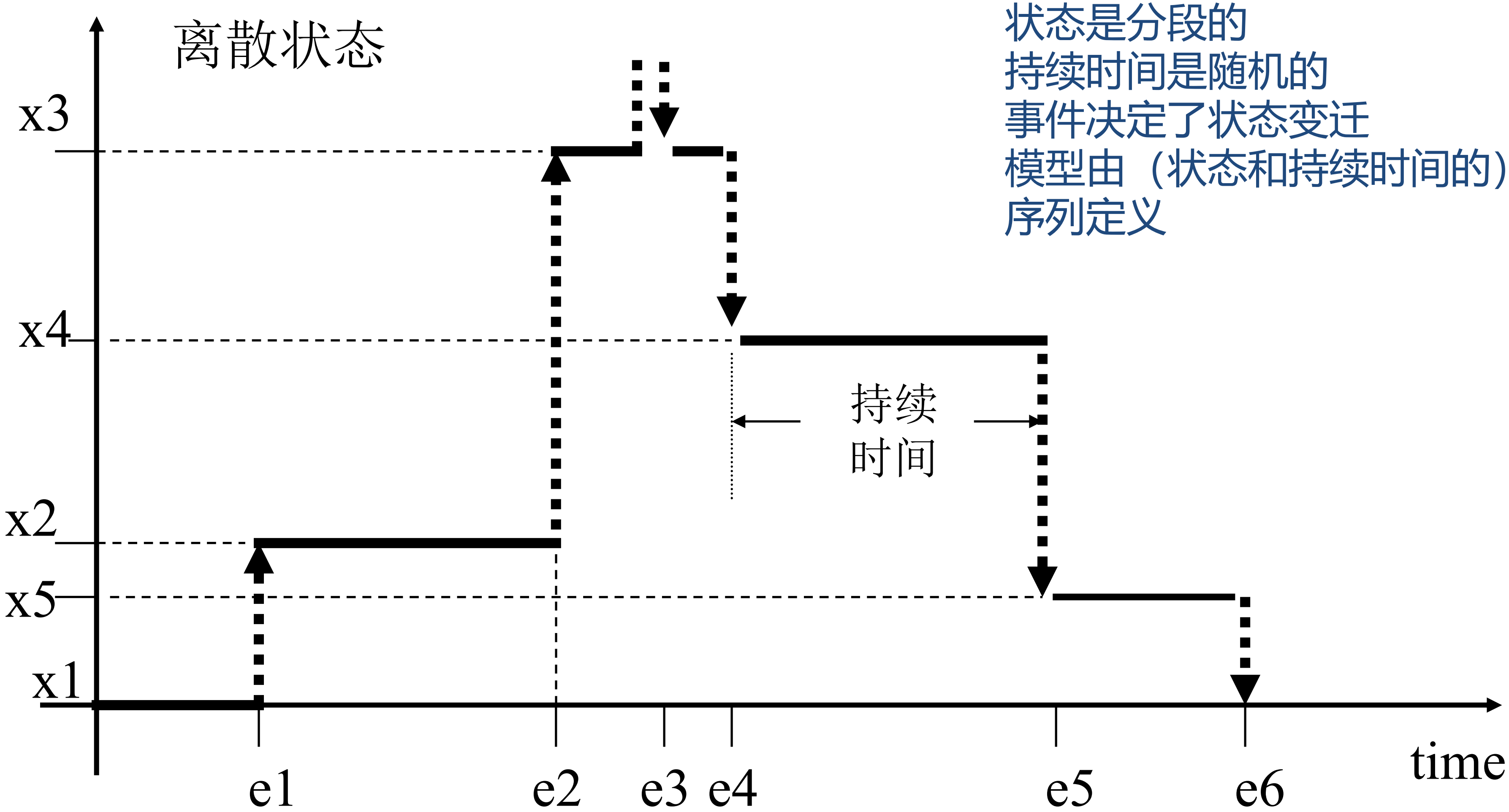


基本概念

什么是 Discrete Event Dynamic Systems (DEDS)?

从1945年-今天

典型的 DEDS



状态是分段的
持续时间是随机的
事件决定了状态变迁
模型由（状态和持续时间的）序列定义

关于 DEDS更多的例子

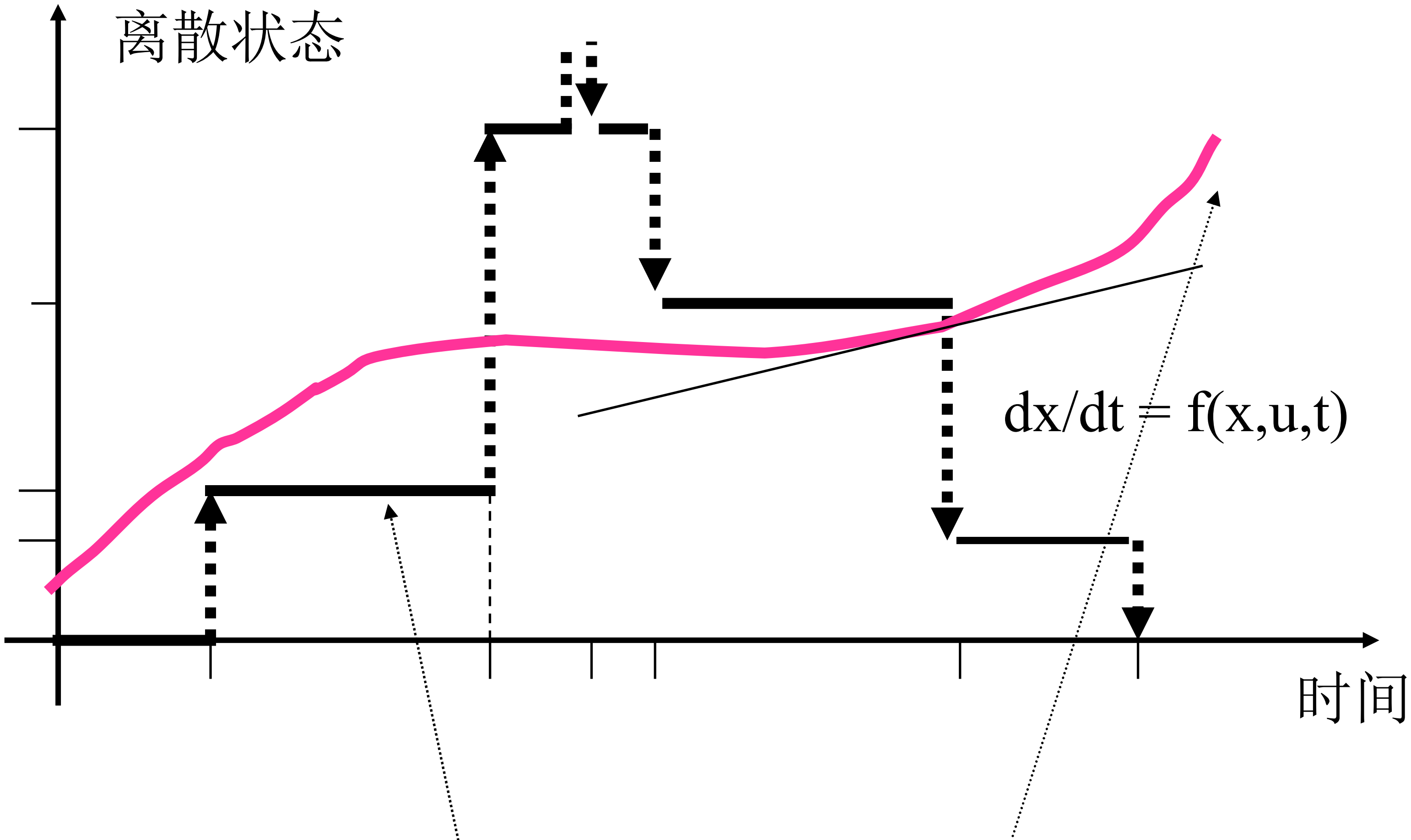
- 自动化制造系统
- 互联网络 - the Internet
- 银行系统
- 交通系统 - Land, Sea, Air

人工系统——随机时间点，随机状态

基本概念

什么是 Hybrid Dynamic Systems (HDS)?

典型的HDS



Hybrid System: 每个状态中都可包含CVDS 行为

发展&历史

CVDS的发展史:

Development of mechanics for CVDS

Self regulating governor for steam engines

<1940

>1940

WWII Servo-mechanism

Modern control theory and practice

DEDS的发展史:

Birth of OR

1945

人工系统出现

1970's

建立理论基础 & 成功实现各种系统

present

DEDS的特点

- DEDS的状态只能在离散时间点上发生变化

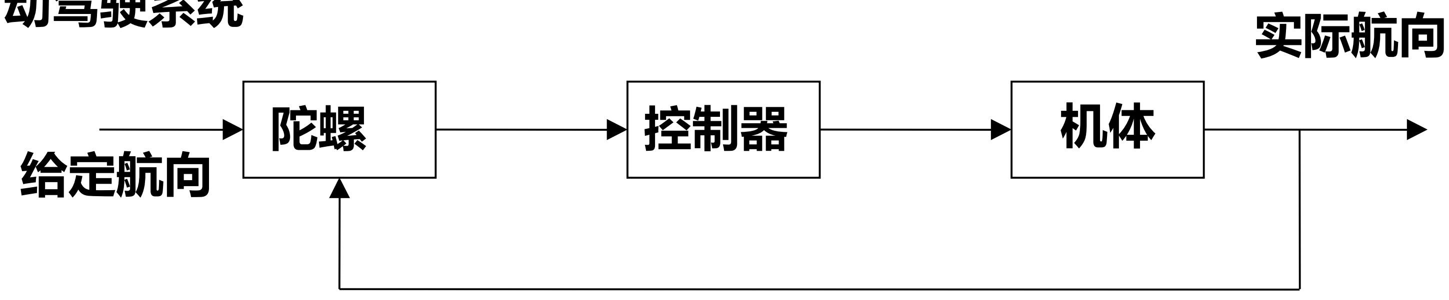
- DEDS的状态变化具有异步性和并发性

- 实际DEDS的状态变化往往呈现出不确定性

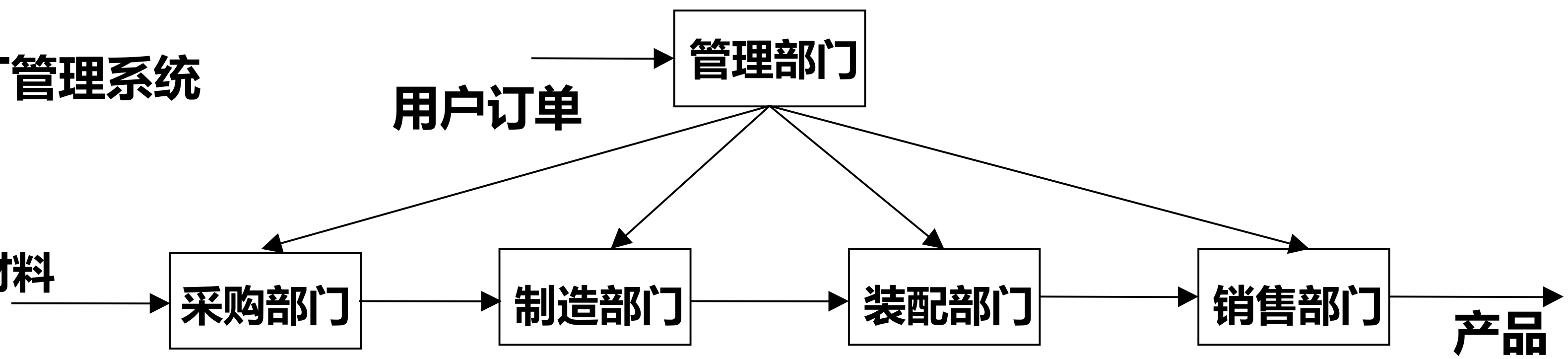
- DEDS服从的是人为的逻辑规则，而不是物理学定律

系统的分类（举例）

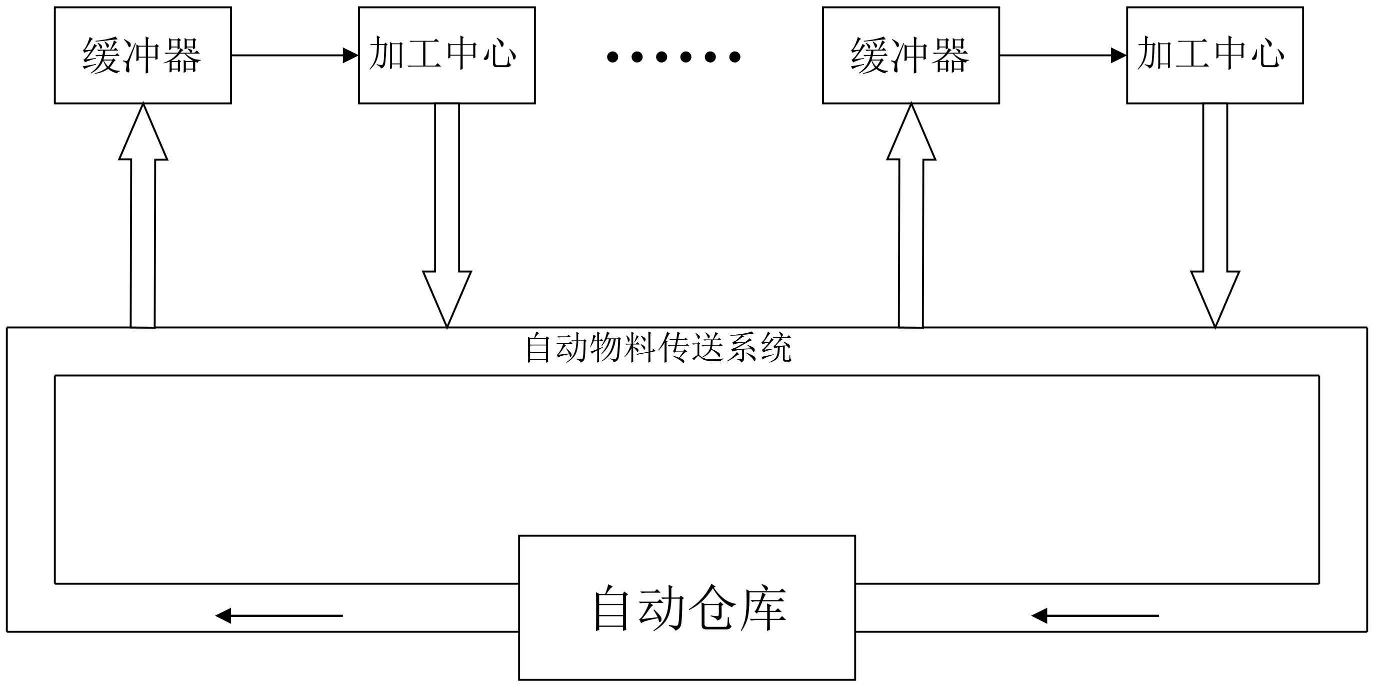
飞机自动驾驶系统



工厂管理系统



系统的分类（举例）



自动生产线

总结

- CVDS（连续变量动态系统）：

  - 连续变量动态系统本质上属于物理世界的范畴。

  - 这类系统的动态过程服从于物理学定律。

- DEDS（离散事件动态系统）：

  - 离散事件动态系统本质上属于人造系统的范畴。

  - 所遵循的是一些复杂的人为规则

基本概念

仿真

为了研究、分析、设计和实现一个系统需要进行实验.

实验的方法：

1）直接在真实系统上进行。

2）先构造模型，然后通过对模型的实验来代替或部分代替）真实系统的实验。

通过模型实验的方法日益被人们所使用：

1）系统处于设计阶段，真实系统尚未建成。

2）在真实系统上实验有风险（发生故障甚至破坏）.

3）在真实系统上实验费用昂贵。

4）多次实验时，难以保证每次实验条件相同。

仿真基础

仿真的定义：

"仿真意指在实际系统尚不存在的情况下，对于系统或活动本质的实现"。（1961）

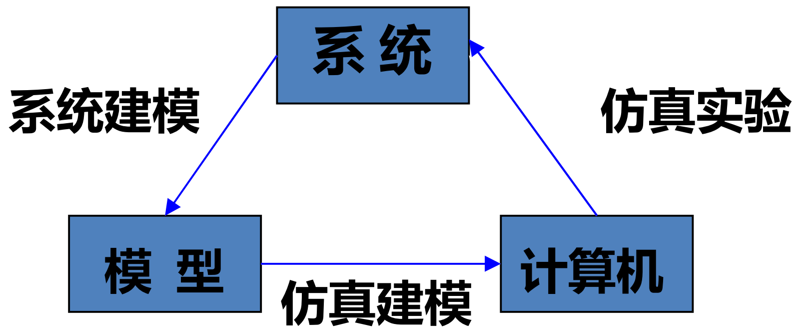
"用能代表所研究的系统的模型作实验"。 （1978）

"所有支持模型建立与模型分析的活动 即为仿真活动"。 （1982）

"仿真是一种基于模型的活动"。 （1984）

"系统,模型,仿真"三者之间有着密切的关系.系统是研究的对象,模型是系统的抽象,仿真是通过对模型的实验以达到研究系统的目的.

系统仿真有三个基本的活动,即系统建模,仿真建模和仿真实验,联系这三个活动的是系统仿真的三要素,即系统,模型和计算机.



仿真的分类

根据模型的种类分类：

物理仿真——按照真实系统的物理性质构造系统的物理模型，并在物理模型上进行实验的过程。

数学仿真——对真实系统进行抽象，用数学关系加以描述而得到系统的数学模型，对数学模型进行实验的过程。（通常是用计算机仿真）

半实物仿真——将数学模型与物理模型甚至实物联合起来进行的过程。

仿真的分类

根据计算机类型分类：

模拟（计算机）仿真 --(并行仿真)
数字（计算机）仿真 --(串行仿真)

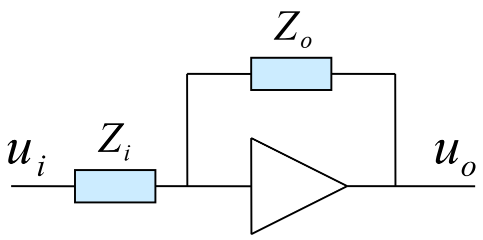
数字-模拟混合仿真

根据仿真时钟与实际时钟的比例关系分类：

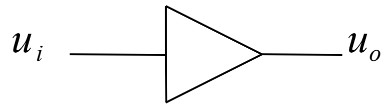
实时仿真 — 仿真时钟与实际时钟完全一致（在线仿真）
(模型仿真速度和实际系统运行速度相同)
亚实时仿真 — 仿真时钟慢于实际时钟（离线仿真）
超实时仿真 — 仿真时钟快于实际时钟（离线仿真）

模拟仿真

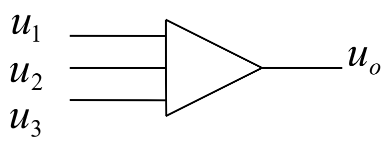
模拟仿真的运算器



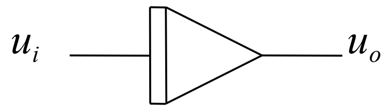
$$u_o = -\frac{Z_o}{Z_i} u_i$$



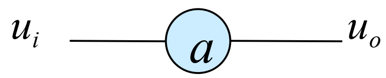
$$u_o = -p u_i$$



$$u_o = -\left(p_1 u_1 + p_2 u_2 + p_3 u_3\right)$$

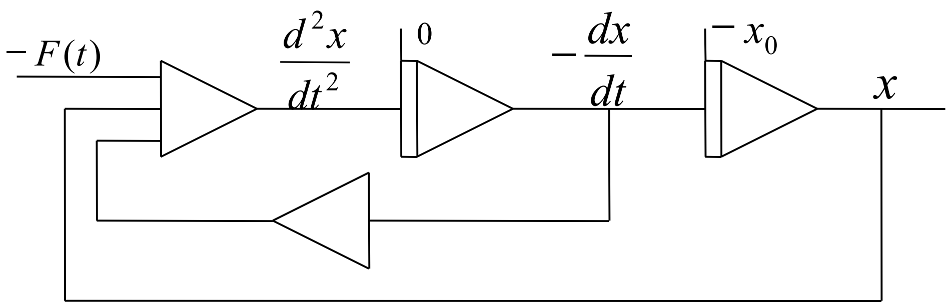


$$u_o = -\frac{1}{T}\int_0^t u_i dt$$



$$u_o = a u_i$$

质量-阻尼-弹簧系统的模拟仿真图



仿真的分类

根据系统模型的特性分类：：

连续（变量动态）系统（ CVDS) 仿真:

—系统状态随时间连续变化的系统 —

1）集中参数模型— 一般用常微分方程（组）描述

(如电路系统，机械动力学系统，生态系统等)

2）分布参数模型— 一般用偏微分方程（组）描述

(如各种物理和工程领域中的“场”问题)

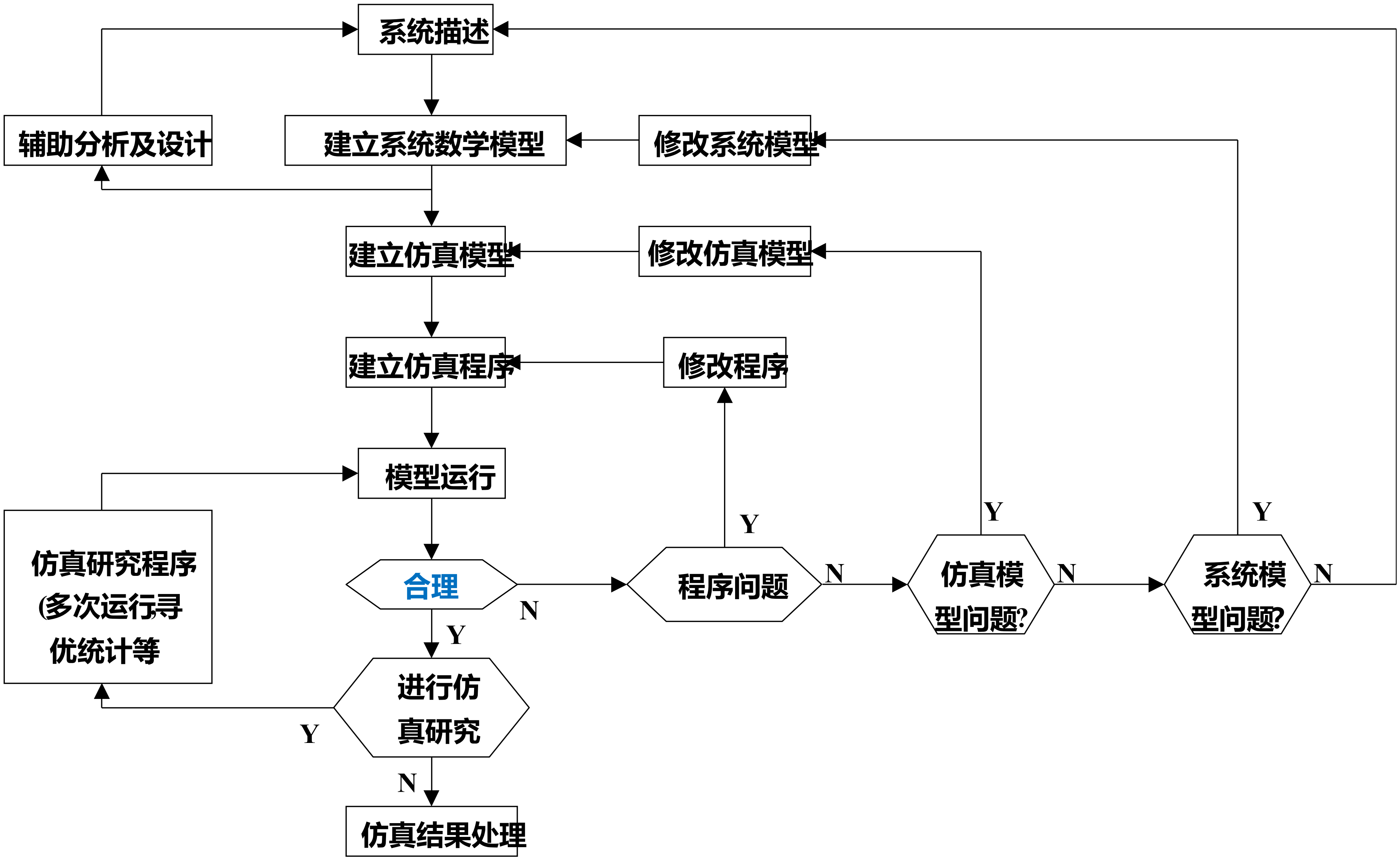
离散（事件动态）系统（DEDS）仿真:

—系统状态在某些离散（随机）时间点上发生（离散）变化的系统—

（ 如交通管理，库存控制，通信系统等 ）

(这类系统的动态特性很难用人们熟悉的数学方程形式加以描述，一般只能借助于流程图或活动图研究的主要目标是系统行为的统计性能而不是行为的点轨迹)

系统仿真的一般步骤



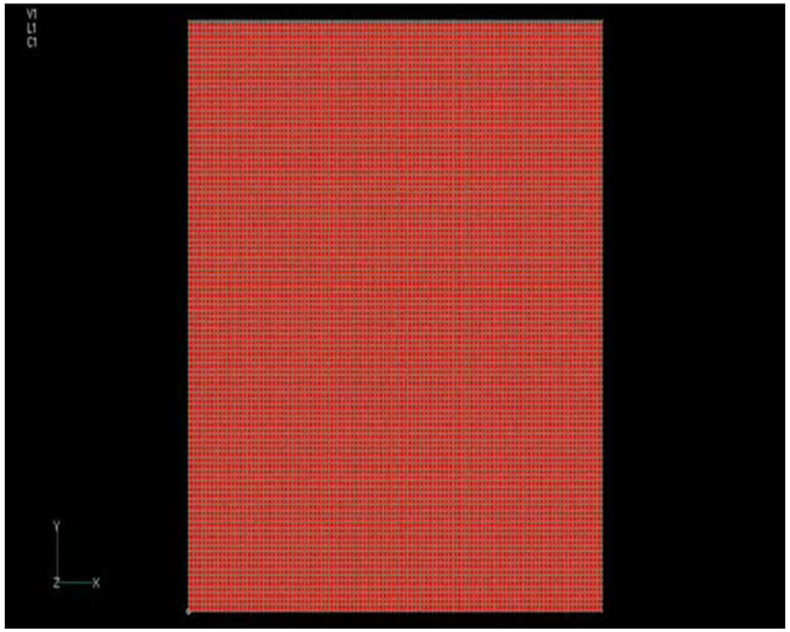
系统描述
辅助分析及设计
建立系统数学模型
修改系统模型
建立仿真模型
修改仿真模型
建立仿真程序
修改程序
模型运行
仿真研究程序
(多次运行寻
优统计等
合理
N
Y
程序问题
Y
N
仿真模
型问题?
Y
N
系统模
型问题?
Y
N
进行仿
真研究
Y
N
仿真结果处理

系统仿真的应用

(1) 系统设计

- 计算机数值仿真方法是现代分析手段中发展最快的方法之一
- 仿真是完善系统模型的主要手段

有限元仿真与优化在结构设计中逐渐普及



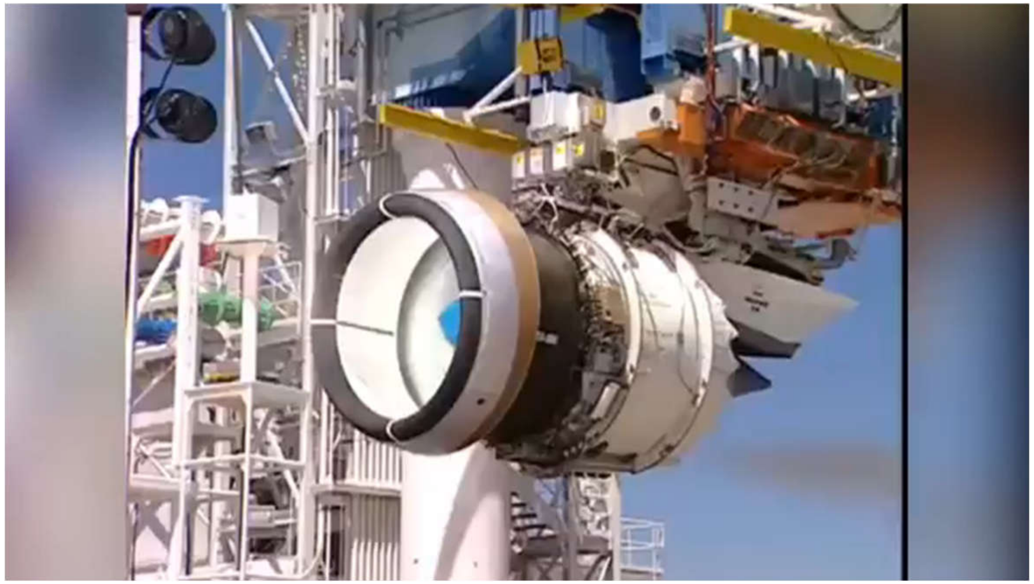
系统仿真的应用

(2) 系统分析

(3) 教育与训练

(4) 产品开发及制造过程

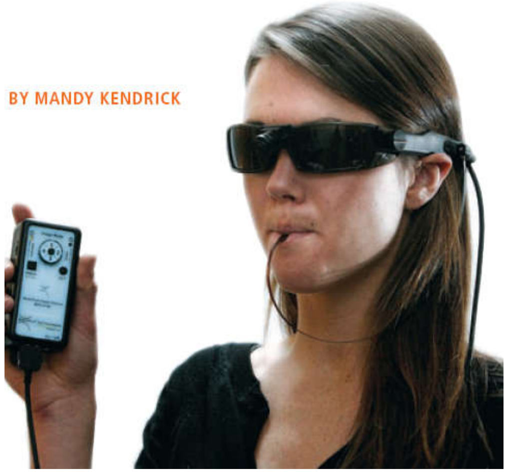
- 虚拟产品开发
- 虚拟制造



(5)产品运维

- 建模方法及工具

建模方法及原则

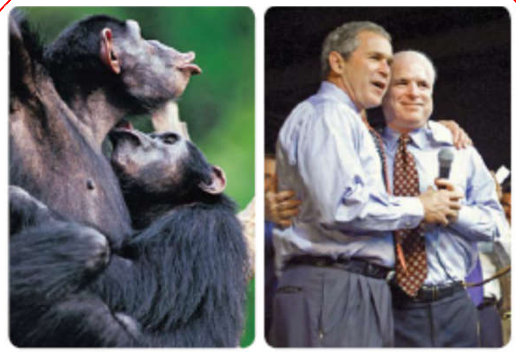


精确反应系统状态、本质特征和变化规律

✓抽象
✓类比
✓移植

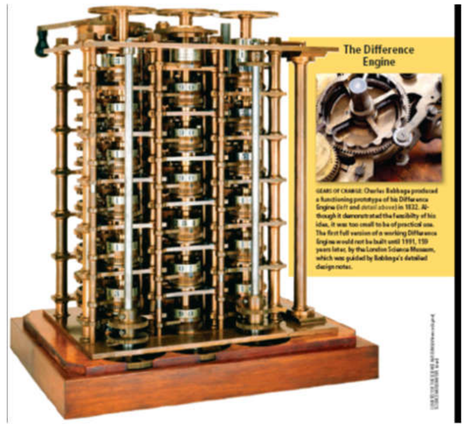
实验统计建模

✓归纳
✓演绎
✓统计



相似性

机理分析建模



综合性方法

Charles Babbage 1848

量纲分析法

- 为了能够应用数学来描述物理对象，我们需要对其定量化。
  物理对象的定量化需要有单位和数值，

▪奥巴马

身高：$1.97663378 \times 10^{-16}$ 光年 **(187cm)**

体重：0.078 吨 **(78Kg)**

▪特朗普

身高：6英尺3英寸 **(190cm)**

体重：239 磅 **(108.4Kg)**

量纲分析法—概念

量纲的定义：

量纲是物理量的单位种类，又称因次。

如长度、宽度、高度、深度、厚度等都可以米、英寸、公尺等不同单位来度量，但它们属于同一单位，即属于同一单位量纲（长度量纲），用[L]表示。

量纲的表示方法：物理量的代表符号外加上中括号。如[L]，[M]，[T]等。

量纲分析法—概念

量纲的分类：基本量纲　导出量纲

基本量纲是一组具有独立性的量纲。在机械领域中有三个基本量纲：[L]，[T]，[M]。

导出量纲由基本量纲组合或推导出来的量纲。如加速度的量纲 $[a]=LT^{-2}$；力的量纲 $[F]=[ma]=MLT^{-2}$

举例1：量纲分析法(自由落体速度)

- 设物体的释放高度为h，自由降落的落地速度为v，不失一般性，我们假定v与物体的质量m、重力加速度g和h有关，即

  $v = f(m, g, h)$

  不做实验，你能求出上述关系式的具体形式吗？

分析…

取基本的量纲为L、T和M，那么有

$[v] = LT^{-1}$， $[m] = M$ ， $[g] = LT^{-2}$， $[h] = L$ 。

容易看出 $[v]$ 的平方等于 $[g]$ 与 $[h]$ 之积，因此 $\frac{v}{\sqrt{gh}} = C(m,g,h)$

- C应该是一个无量纲的数，即上式中右边的变量应该组合成为无单位的纯数。由于M在除m外的其他变量中都未出现，所以m不能组合成无量纲的纯数，应该右边的变量中去掉。同理可知变量 g, h 也应该从右边的变量中去掉。这表明C应该是一个无量纲的纯常数，这样我们就得到 $v = C\sqrt{gh}$

其中常数C可以通过实验测定出来。

量纲分析法

量纲和谐原理

凡是正确反映客观规律的物理方程，其各项的量纲都必须是一致的，即只有方程两边量纲相同，方程才能成立。

- **a.一个方程在量纲上应是和谐的，所以可用来检验经验公式的正确性和完整性。**
- **b.量纲和谐原理可用来确定公式中物理量的指数。**
- **c.可用来建立物理方程式的结构形式。**

量纲分析法

雷利法

解题步骤：首先找出影响系统的物理量，并用它们写出假拟的**指数方程**；

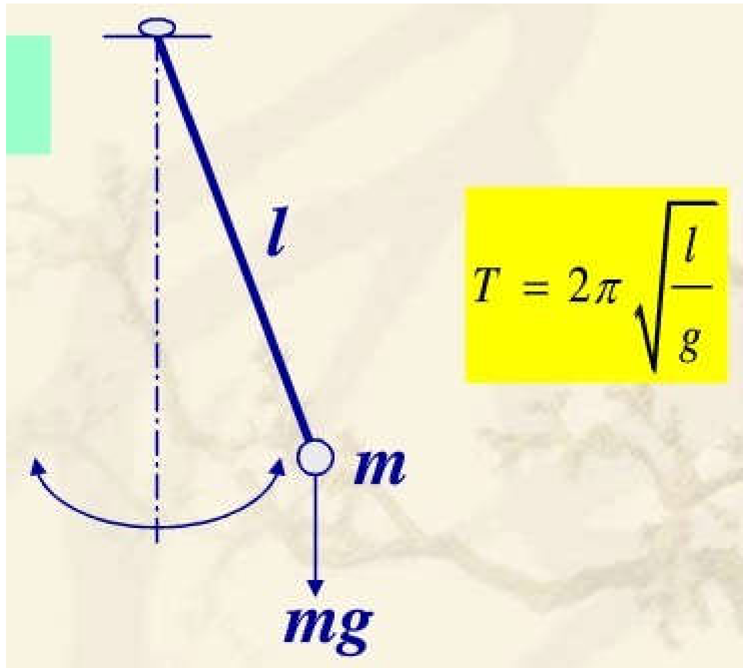
然后以对应的量纲代替方程中的物理量本身，并根据量纲和谐性原理求出各物理量的指数，整理出最后形式。

π定理

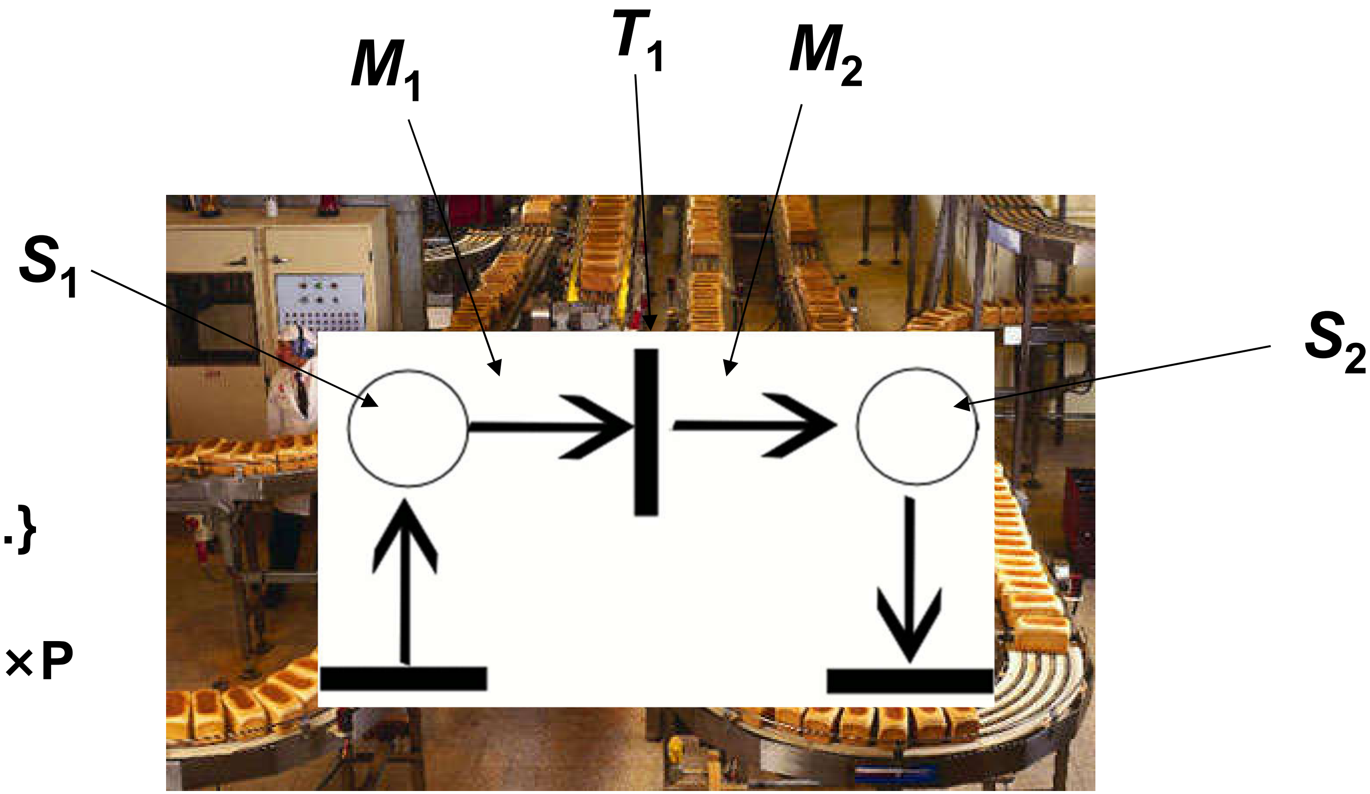
- **对于某个物理现象或过程，如果存在有n个变量互为函数关系，**
- $$f(a1,a2,\ldots an)=0$$
- **而这些变量含有m个基本量纲，可把这n个变量转换成为有(n-m)=i个无量纲量的函数关系式**
- $$F(\pi 1,\pi 2,\ldots \pi n-m)=0$$
- **这样可以表达出物理方程的明确的量间关系，并把方程中的变量数减少了m个，更为概括集中表示物理过程或物理现象的内在关系。**

举例2：量纲分析法（单摆的周期）

- 质量*m*的小球系在长度为*l* 的线的一端，偏离平衡位置后小球在重力*mg*的作用下做往复摆动，求小球的摆动周期*T*。



抽象方法



S={S_1,S_2…}

T={{S_1,S_2}…}

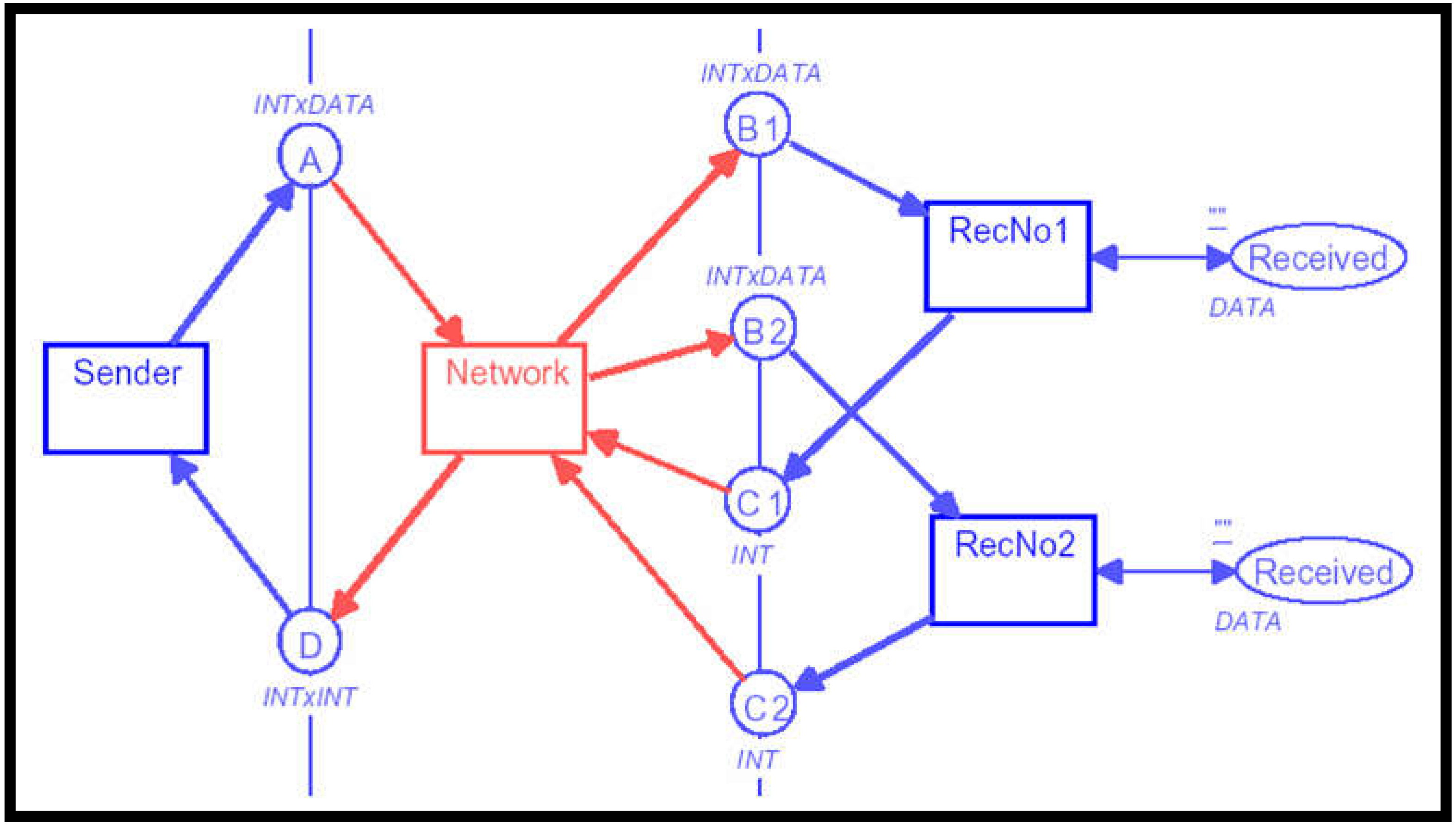
M: P×T ∪ T×P

如何用数学语言描述包装过程

生产过程模型N={S,T,M}

生产过程描述： N_0, N_1,N_2,N_3…

更复杂的情形…


INTxDATA
A
Sender
Network
D
INTxINT
INTxDATA
B1
INTxDATA
B2
C1
INT
C2
INT
RecNo1
RecNo2
""
Received
DATA
""
Received
DATA

基本工具

概率: 离散随机变量*A*取*a*值的概率为*P*(*A*=*a*) $\in (0,1)$，常写作*P*(*a*)

- 归一化 $\Sigma_{all\ a} P(A=a)=1$
- 联合概率 $P(A=a,B=b)=P(a,b)$
- 边缘化 $P(A=a)=\Sigma_{all\ b} P(a=a,B=b)$
- 条件概率 $P(a|b)=\frac{P(a,b)}{P(b)}$ *b*发生时，*a*发生的概率
- 联合概率

$$P(a,b)=P(a)P(b|a)=P(b)P(a|b)$$

基本工具

概率: 离散随机变量*A*取*a*值的概率为*P*(*A*=*a*) $\in (0,1)$，常写作*P*(*a*)

▪贝叶斯规则 $P(a|b) = \frac{P(b|a)P(a)}{P(b)}$ $P(a|b,C) = \frac{P(b|a,C)P(a|C)}{P(b|C)}$

系统变量 θ 观测变量 D

似然函数，$\int p(D|\theta)d\theta \neq 1)$

先验概率

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$$

后验概率

证据因子

▪相互独立 $P(a,b) = P(a)P(b)$ $P(a|b) = P(a)$ $P(b|a) = P(b)$

*Iff A*与*B*相互独立

基本工具

概率: 连续变量x 使用概率密度函数 $p(x) \in [0, \infty]$

- 概率质量 $P(x_1 < X < x_2) = \int_{x_1}^{x_2} p(x)dx \in [0, \infty]$

$$\int_{-\infty}^{\infty} p(x)dx = 1 \qquad p(x) = \int_{-\infty}^{\infty} p(x,y)dy$$

- 期望 函数f在概率分布P下的期望定义为

离散变量 $$\mathbb{E}_P|f| = \sum_a P(a)f(a)$$

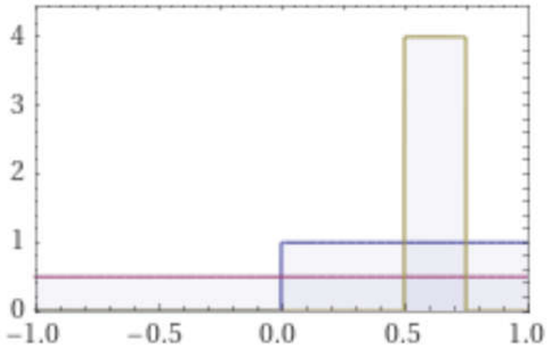
连续变量 $$\mathbb{E}_P|f| = \int_x p(x)f(x)dx$$

- 方差 $V(f) = \mathbb{E}[f(x) - \mathbb{E}[f(x)]])^2 = \mathbb{E}[f(x)^2] - \mathbb{E}[f(x)]^2$
- 标准差 $std(f) = \sqrt{V(f)}$
- 协方差 $Cov(x,y) = \mathbb{E}_{x,y}[(x - \mathbb{E}(x))(x - \mathbb{E}(x))]$

基本工具

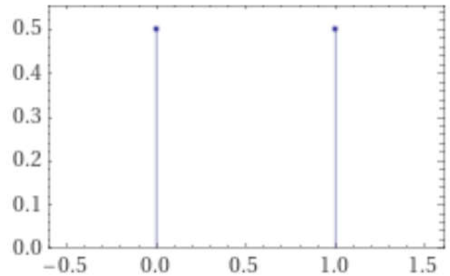
分布:

▪均一分布(Uniform dis.)



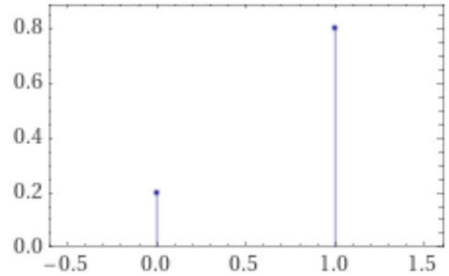
$$P(A=a_i)=1/K, i=1,\ldots,K$$

▪伯努利分布（Bernoulli dis.）



$\mu=0.5$

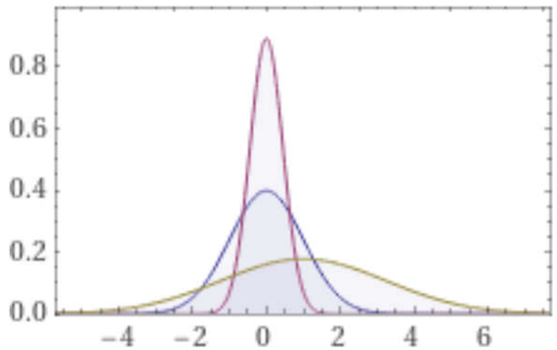
$$P(x|\mu)=\mu^x(1-\mu)^{1-x}$$



$\mu=0.8$

$$P(x|\mu)=\mu^x(1-\mu)^{1-x}$$

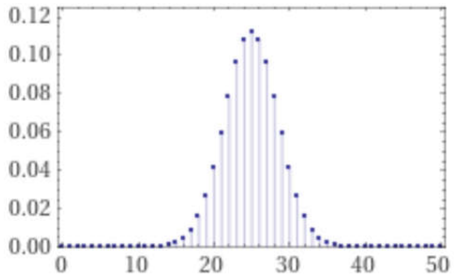
▪正态分布（Normal dis.）



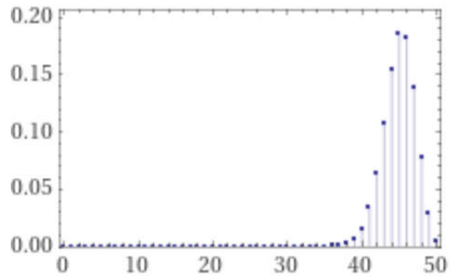
单变量 $p(x|\mu,\sigma^2)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$

D维变量 $p(x|\mu,\Sigma)=(2\pi)^{-\frac{D}{2}}|\Sigma|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}(x-\mu)^\top\Sigma^{-1}(x-\mu)\right)$

▪二项分布（Binomial dis.）



$\mu=0.5 \quad n=50$



$\mu=0.9 \quad n=50$

$$P(m|N,\mu)=\binom{N}{m}\mu^m(1-\mu)^{N-m}$$

这里 $\binom{N}{m}=\frac{N!}{(N-m)!m!}$

基本工具

线性代数:

- 计算

$$\underset{n\times m}{A}\cdot\underset{m\times p}{B}=\underset{n\times p}{C}\qquad C_{ij}=\sum_{k=1}^{m}A_{ik}B_{kj}$$

$$(AB)C=A(BC)$$
$$A(B+C)=AB+AC$$
$$(A+B)C=AC+AB$$
$$(A+B)^T=A^T+B^T$$
$$(AB)^T=B^TA^T$$
$$AB\neq BA$$

- 属性

对角阵 $A_{ij}=0,\forall i\neq j$

单位阵 $I_{ii}=1\quad AI=IA=A$

$$AA^{-1}=A^{-1}A=I$$
$$(AB)^{-1}=B^{-1}A^{-1}$$
$$(A^\top)^{-1}=(A^{-1})^\top$$

迹 $\mathrm{Tr}(A)=\sum_i A_{ii}$

行列式 $|A|\qquad |a|=a$

$$|AB|=|A||B|$$
$$|aA|=a^n|A|$$
$$|A^{-1}|=1/|A|,\ |A|\neq 0$$

$|A|=0\quad \underset{n\times n}{A}\longrightarrow\underset{r\times r}{A}$

奇异阵　秩r

特征值 $\lambda_i\qquad Au_i=\lambda_i u_i\qquad u_i^\top u_j=I_{ij}$

- 范式

0-norm: $\|x\|_0=$ 非0单元个数

2-norm: $\|x\|_2=\left(\sum_{i=1}^n x_i^2\right)^{1/2}$

1-norm: $\|x\|_1=\sum_{i=1}^n|x_i|$

∞-norm: $\|x\|_\infty=\max_{i=1}^n|x_i|$

基本工具

微积分: $f'(x) = \frac{df}{dx} = \lim_{\delta \to 0} \frac{f(x+\delta)-f(x)}{\delta}$ $\quad f''(x) = \frac{d^2 f}{dx^2} = \frac{df'}{dx}$

- 常用 $c' = 0,\ (cx)' = c,\ (cx^a)' = cax^{a-1}$ $\quad (\log x)' = 1/x,\ (e^x)' = e^x$

  $(f(x)+g(x))' = f'(x)+g'(x)$ $\quad (f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$

- 链式法则 $$\frac{df(y)}{dx} = \frac{df(y)}{dy}\frac{y}{dx}$$

- $f(x_1, \ldots, x_n)$ $$\frac{\partial f}{\partial x_i} = \lim_{\delta \to 0} \frac{f(x_1, \ldots, x_i + \delta, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_i, x_{i+1}, \ldots, x_n)}{\delta}$$

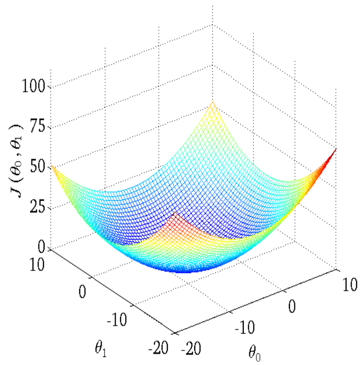
基本工具

微积分: $f'(x) = \frac{df}{dx} = \lim_{\delta\to 0}\frac{f(x+\delta)-f(x)}{\delta}$ $\qquad$ $f''(x) = \frac{d^2 f}{dx^2} = \frac{df'}{dx}$

- $f(x_1,\ldots,x_n)$ $\qquad$ $\frac{\partial f}{\partial x_i} = \lim_{\delta\to 0}\frac{f(x_1,\ldots,x_i+\delta,x_{i+1},\ldots,x_n)-f(x_1,\ldots,x_i,x_{i+1},\ldots,x_n)}{\delta}$

- 梯度 $\qquad$ $\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$

$$x = (x_1,\ldots,x_n)$$



$$\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

一阶导：指向变化最快方向的向量 $\qquad$ 二阶导：n×n 海森矩阵

- 判定函数局部极值的充分必要条件

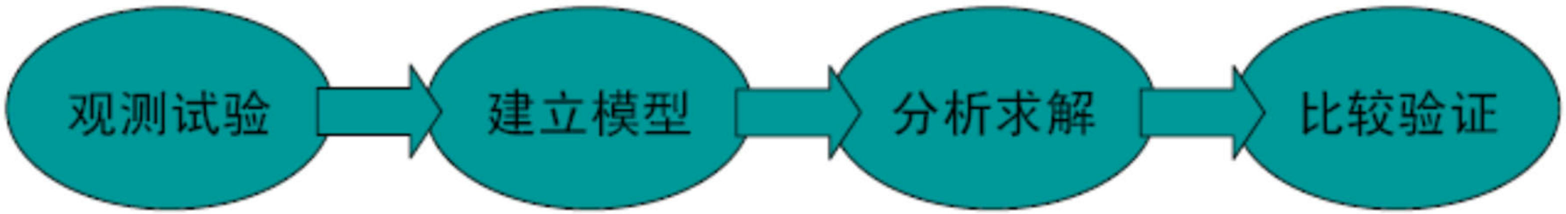
$\nabla f(x) = 0$ $\qquad$ $\nabla^2 f(x)$ 是半正定矩阵

- 凸函数

$$\forall x, y, \lambda \in [0,1],\ f(\lambda x + (1-\lambda) y) \le \lambda f(x) + (1-\lambda) f(y)$$

总结

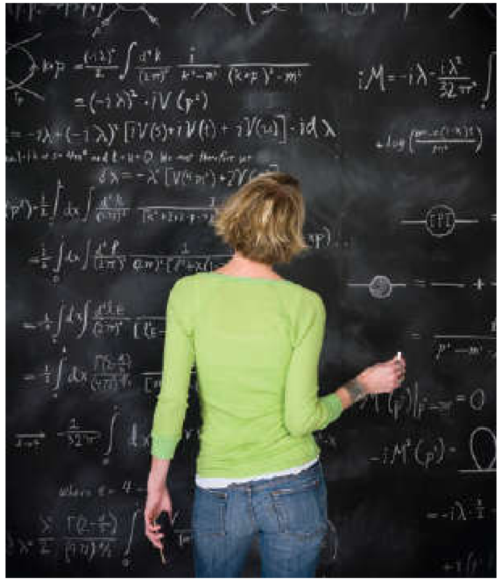
- 建模的步骤



解析方法、实验方法、数值方法相互配合/补充

建模所要求的能力

- 分析综合能力
- 抽象概括能力
- 想象洞察能力
- 运用数学工具的能力
- 运用工程软件的能力



相关数学工具（一种或几种）

Matlab（Octave） **矩阵运算**

Mathematica

Mape **符号计算系统**

SPSS/SAS **数据库与统计**

Ansys/Abacus/ALGOR **FEA工具**

需要掌握编程语言（一种或几种）

C++

C#

快速建立原型

Python

AutoLisp/Grip

习题

设 $p(\mathbf{x}|\Sigma) \sim N(\mu, \Sigma)$ 这里 μ 已知，Σ 未知。如果 Σ 的最大似然估计为 $\widehat{\Sigma} = \frac{1}{n}\sum_{k=1}^{n}(x_k - \mu)(x_k - \mu)^t$，证明以下论述：

1) 证明矩阵等式$a^t A a = tr[Aaa^t]$，这里矩阵的迹$tr[A]$表示$n \times n$维矩阵A的对角线元素之和，a为一个向量。

2) 证明似然函数可以写为以下形式：

$$p(\mathbf{x_1}, ..., \mathbf{x_n}|\Sigma) = \frac{1}{(2\pi)^{nd/2}}|\Sigma^{-1}|^{n/2} exp\left[-\frac{1}{2}tr\left[\Sigma^{-1}\sum_{k=1}^{n}(\mathbf{x_k} - \mu)(\mathbf{x_k} - \mu)^t\right]\right]$$

注：多维正态分布$p(\mathbf{x}|\Sigma) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} exp\left[-\frac{1}{2}\left\{(\mathbf{x_i} - \mu)^T \Sigma^{-1}(\mathbf{x_i} - \mu)\right\}\right]$