信息论

信号传输与处理的理论基础

凸性概念及应用

凸性(Convexity)

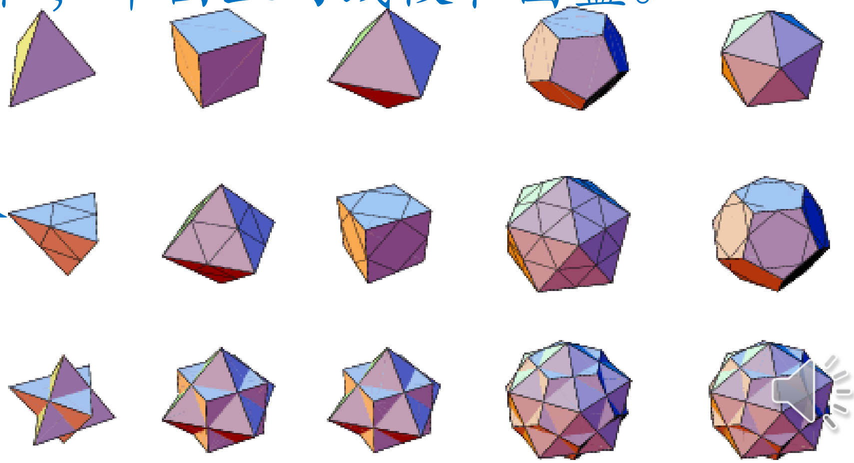
* 凸性在网络通信、计算机控制、人工智能、组合优化、数理经济学（如博弈论）、金融风险分析等领域具有越来越多的关键应用，也是理解信息论中很多重要概念规律的基本工具之一。

* (1) 几何凸性 Geometric Convexity
* (2) 函数的凸性 Function Convexity

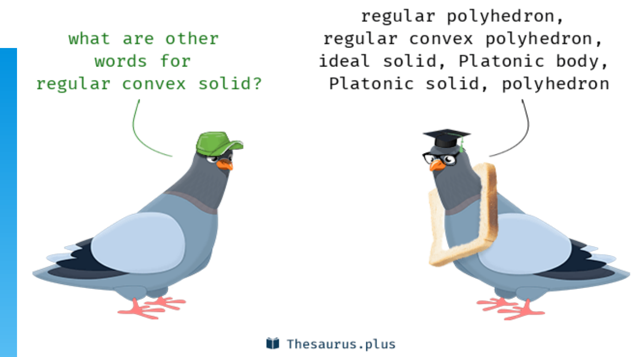
* 教程中在很多地方用到凸性的概念和方法，本讲义在此将其集中阐述，以便于应用和查阅。

几何凸性

* （1）几何定义：n维空间中的一个集合Ω，如果其中任取两点，连接该两点的线段总是完全落在Ω内，则Ω定义做是一个凸集。
* （2）解析定义：n维空间中的一个集合Ω，如果对其中任取两点x和y，以及任意的实数$0\le t\le 1$，点$tx+(1-t)y$仍在Ω内，则Ω定义做是一个凸集。
* 【注】以上两个定义是等价的。
* 【注】是否存在“凹”集合的定义？没有！
* （3）典型的例子：
* 三维空间中的平面、球体、圆柱体；平面上的线段和圆盘。
* (4) 习题：
* 以下是凸体的例子吗?
* 平面上的圆环、球的表面、中空的立方体
* 右图中的哪些实体是凸的?



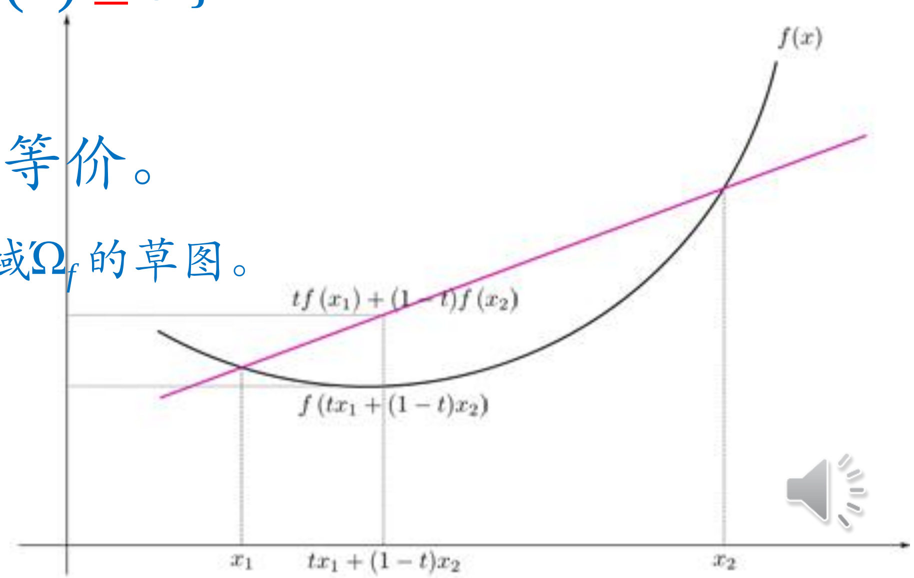
函数的凸性



* （1）解析定义：n元函数$f(x)$定义做凸函数，是指对定义域中的任何两点x和y以及任意的$0\le t\le 1$，恒满足以下不等式：
* $$f(tx+(1-t)y) \le t\, f(x) + (1-t)\, f(y)$$
* （2）几何定义：
* n元函数$f(x)$定义做凸函数，是指其相关的区域
* $$\Omega_f = \{(\mathbf{x},s): f(x) \le s\}$$
* 是n+1维空间中的凸集合。

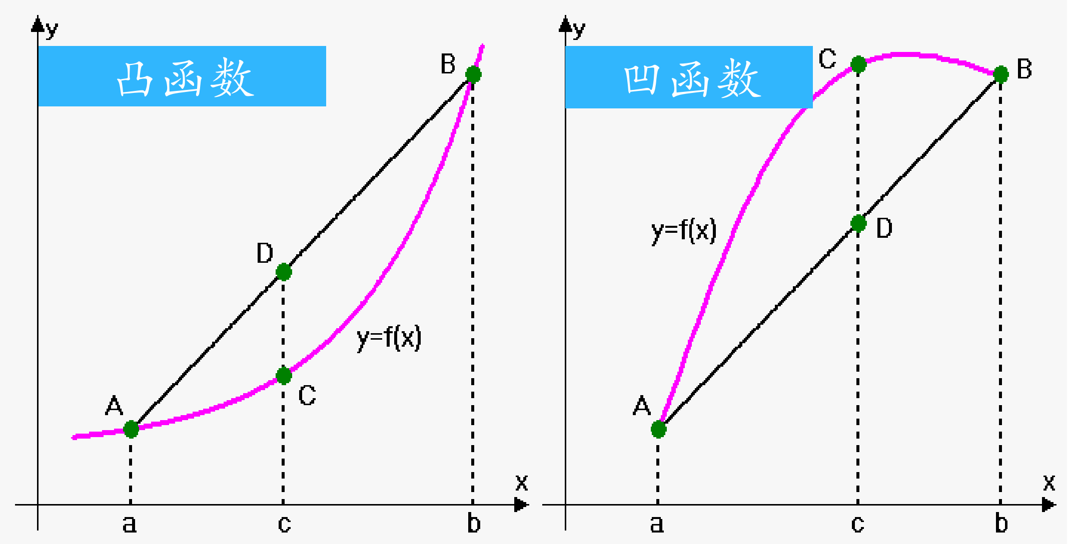
（3）习题：验证上述两个定义等价。

提示：画出函数$f(x)$所决定的的区域Ω_f的草图。



函数的凸性

* （1）解析定义：n元函数$f(x)$定义做凹函数，是指对定义域中的任何两点x和y以及任意的$0\le t\le 1$，恒满足以下不等式：
* $$f(tx+(1-t)y) \geq t\,f(x) + (1-t)\,f(y)$$
* （2）几何定义：
* n元函数$f(x)$定义做凸函数，是指其相关的区域
* $$\Omega_f = \{(\mathbf{x},s) : f(x) \geq s\}$$
* 是n+1维空间中的凸集合。



函数的凸性

- 典型实例
- $a_1x_1+\ldots+a_nx_n$、$exp(x)$、$-logx$、$xlogx$、$x^m(m\geq 1)$是凸函数。
- 基本性质：
- （1） 函数$f(x)$凸， 当且仅当函数 $-f(x)$凹；
- （2） 若函数$f(x)$和$g(x)$凸， 则$f(x)+g(x)$凸；
- （3） 若函数$f(x)$凸， 则集合$\{x: f(x)\leq 0\}$是凸集合；
- （4） 若函数$f(y)$和$g(x)$凸， 则$f(g(x))$凸；
- (5) 若函数$f(y)$和$g(x)$凸， 则$max(f(x),g(x))$凸；
- 习题： 根据凸函数的定义检验上述性质。
- 注1： 函数并非是非凸即凹，实际上很多函数既不是凸函数、也不是凹函数，例如$sinx$、$tanx$等。
- 注2：某些函数既不凸、也不凹，但在某些区域上是凸函数、另外的区域上是凹函数，即所谓局部凸函数。
- 你能指出$\sin x$在$[0,2\pi]$的哪个子区间上是凸函数、那个子区间上是凹函数吗?

函数的凸性

* 函数的凸性判定准则
* （1）二阶可微分的n元函数$f(x)$在区域G上是凸函数，
* 当且仅当 对称矩阵$[\partial^2 f/\partial x_i \partial x_j]$是正定矩阵，即当x在G
* 上任意变化、参数$u_1,\ldots,u_n$任意变化时，恒有不等式
* $$\sum_{i,j=1}^{n} u_i u_j \partial^2 f/\partial x_i \partial x_j \geq 0.$$

（2）特例：一元二阶可微分的函数$f(x)$在区域G上是凸函数，当且仅当x在G上变化时恒成立不等式$f^{(2)}(x) \geq 0$。

习题1：从普遍的准则（1）导出准则（2）。

习题2：运用上述准则重做上一页的习题，包括验证哪些典型的实例，如xlogx凸等。但该种检验不适合于上页的性质（5），即函数$max(f(x),g(x))$，为什么？

习题3：证明普遍的准则（1）。提示：运用展开到二阶项的精确的Taylor公式，该公式具有带二重积分所表达的误差项，然后根据凸函数的定义进行分析。

习题4：请写出关于凹函数的普遍判定准则、针对一元函数的特殊准则。

函数的凸性

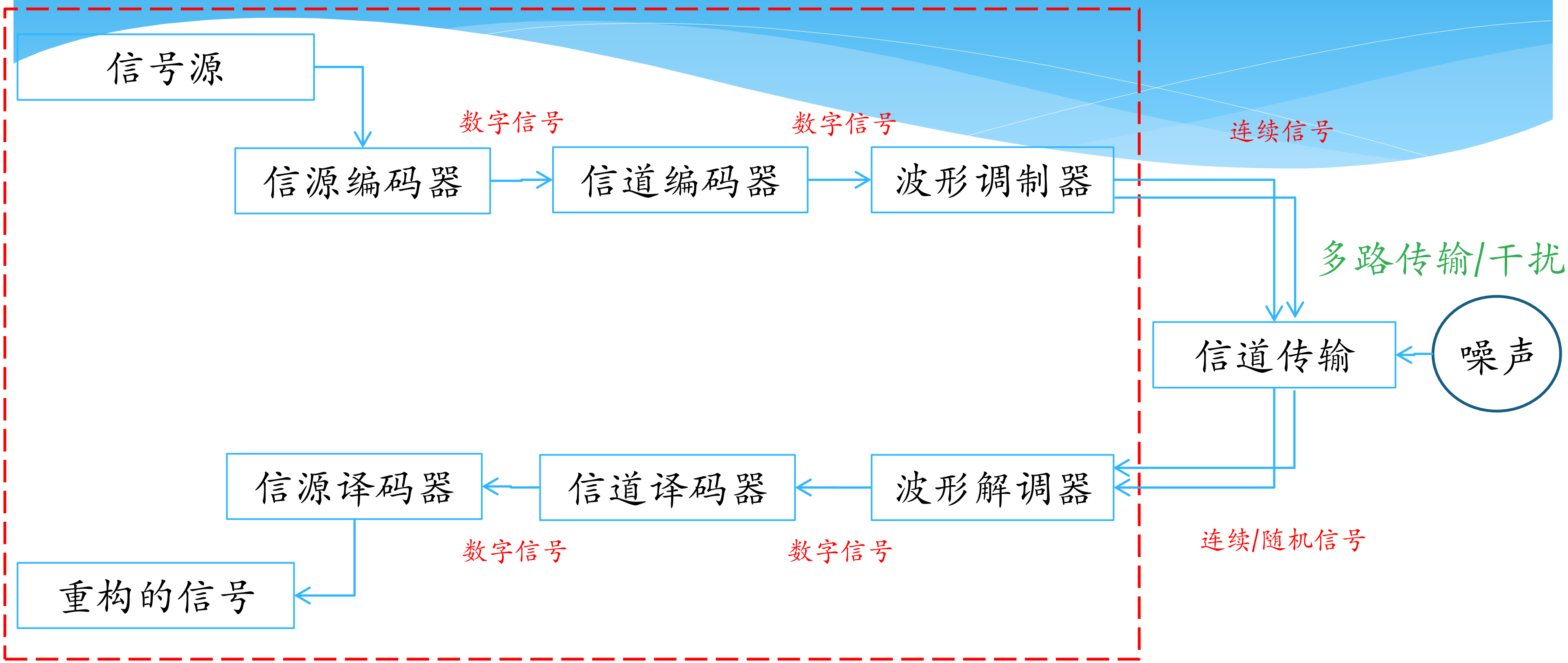
* 更多的一些实例（习题）
* 运用前述判定准则，判定以下x的函数哪些是凸函数、
* 哪些是凹函数、哪些既不是凸的也不是凹的?
* （1）Gauss函数 $\exp(x^TAx)$,x是n维向量，A是 nXn阶正定矩阵；如果A是负定矩阵呢?
* （2）函数 $\exp(x^TAx+b^Tx)$, x、b是n维向量，A是正定矩阵 ；
* （3）函数$\log(x^TAx)$, x和A同问题（1）；
* （4）$(x^TAx)^m$, x和A同问题（1），指数m是大于1的实数。

函数的凸性

* 一、信息论和数字通信领域中的大量重要的表达式是凸函数
* 或凹函数，例如：
* 信息熵H[X]：X的概率分布的凸函数；
* 互信息量I(X;Y): x的概率分布的凸函数、条件概率的凹函数；
* 二、凸性的典型应用
* 在网络工程和自动控制、模式识别、信号处理等领域，
* 存在大量形如 $\min f(x)$ s.t. $g_1(x) \leqq a_1, g_m(x) \leqq a_m$ 的带凸约束的优化问题，其中$f(x)$和诸$g_i(x)$均为凸函数。对这类问题，能够建立具有多项式时间复杂度的算法实施快速求解。

凸性是导致上述各种优化算法得以建立的实时性特点！

课程概要



* 信号传输与处理的概念模型