智能优化算法

课程要求：

1. 课程考核

① 考勤，随堂练习题(准备纸、笔)，30%

② 期末考试待定，40%

③ 程序实现作业，抄袭，迟交扣分，30%

2. 联系方式

① zhanghy307@163.com

② 手机：15122929930

③ 有事请假，不得无故缺课

3. 课程推荐教材:

1.**智能优化算法及其matlab实例，包子阳，余继周等，电子工业出版社**

2. 凸优化， 王书宁，黄晓霖等译，清华大学出版社.

3. matlab智能算法， 温正，孙华克等，清华大学出版社

4. Numerical optimization, Jorge Nocedal , Stephen J. Wright, springer

4. 所需计算机语言与软件

① MATLAB为主

② Visual Studio/opencv

③ C或者C++/python

相关基础知识：

- 线性代数
- 数学分析
- 数字图像处理基础
- matlab编程能力

本课程涉及内容：

传统优化算法

- 凸函数与凸优化问题
- 主对偶问题
- 增广拉格朗日算法及应用

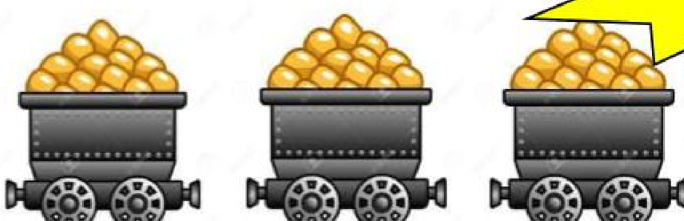
智能算法

- 遗传算法，免疫算法
- 蚁群算法，粒子群算法
- 模拟退火算法，禁忌搜索算法
- 人工神经网络与机器学习算法
- 小波分析算法

1.1 人工智能：大数据时代

大数据 ≠ 大价值

机器学习



有效的数据分析

- 无处不在的数据

(google, 百度，QQ,微信，卫星，监控...)

- 数据类型

字符、数字、图表、图像、视频、点击流..

- 数据处理分析现状

数据量大，复杂多样，人工处理困难

- 计算机技术的发展

使用计算机代替人工分析处理数据，从中抽取有用的信息

1.2 人工智能

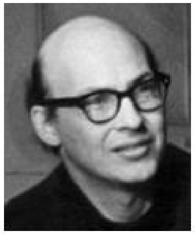
- 人工智能：
  - Artificial Intelligence： AI
  - **1956**年，几个计算机科学家相聚在**达特茅斯**会议，提出了“人工智能”的概念：
  - **用计算机来构造复杂的、拥有与人类智慧同样本质特性的机器**
  - 是对人的意识、思维的信息过程的模拟。能像人那样思考、也可能超过人的智能。

Artificial Intelligence (AI), 1956 -

1956年夏　美国达特茅斯学院

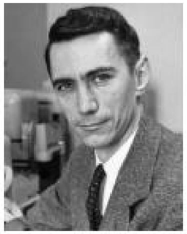
J. McCarthy

"人工智能之父"
图灵奖(1971)



M. Minsky

图灵奖(1969)



C. Shannon

"信息论之父"

H. A. Simon

图灵奖(1975)
诺贝尔经济学奖(1978)

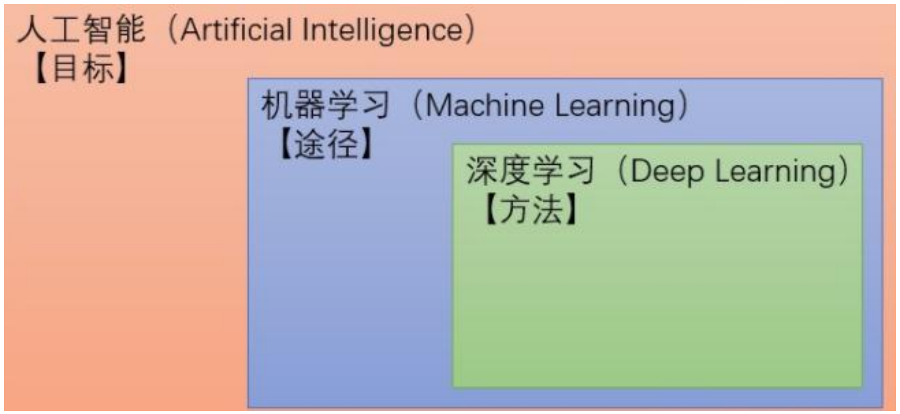
A. Newell

图灵奖(1975)

……

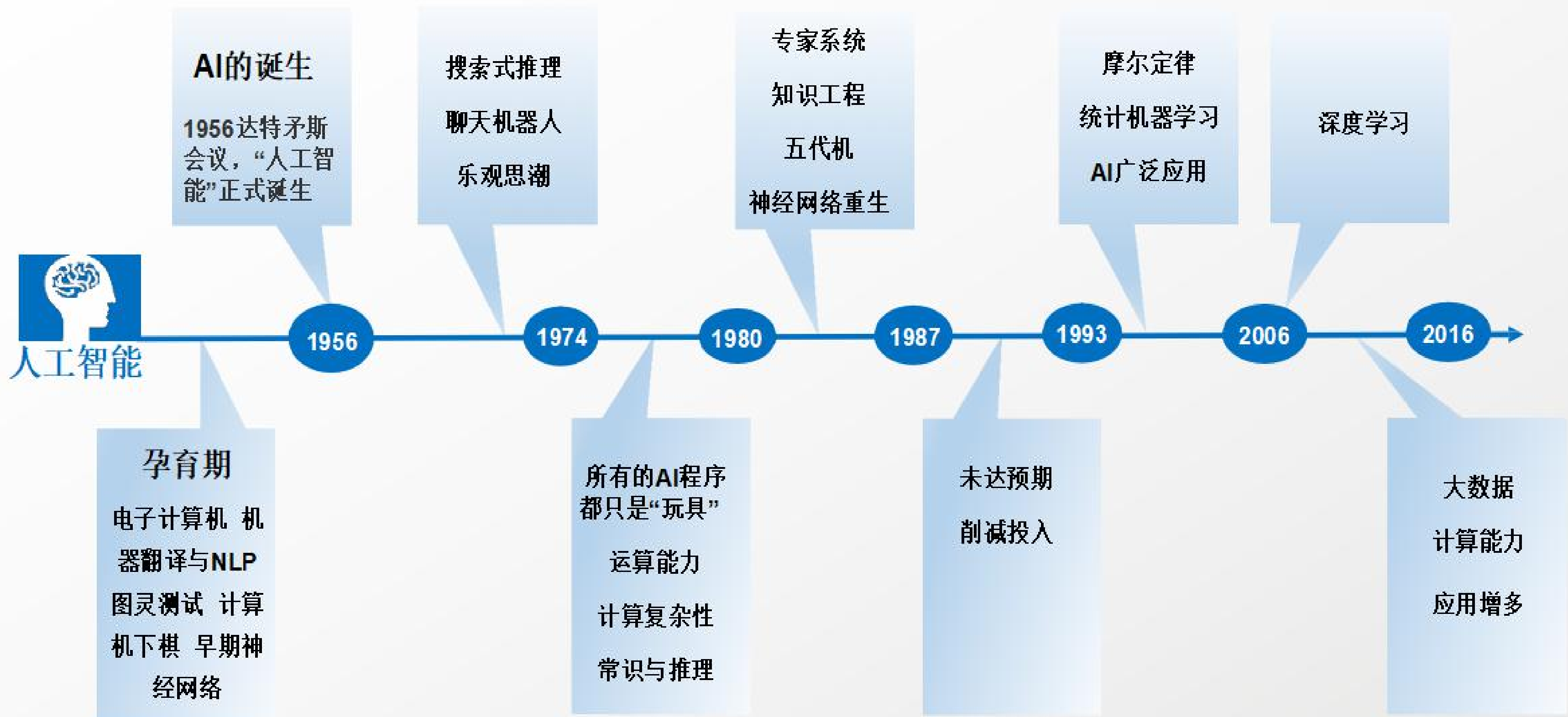
……

达特茅斯会议标志着人工智能这一学科的诞生



1. 人工智能：是终极目标，**希望机器有足够的智能以服务于人类**
2. 机器学习：是实现人工智能的一种主要途径
3. 深度学习：是机器学习中重要的一种方法，可以应对更加复杂的问题。

人工智能发展历程

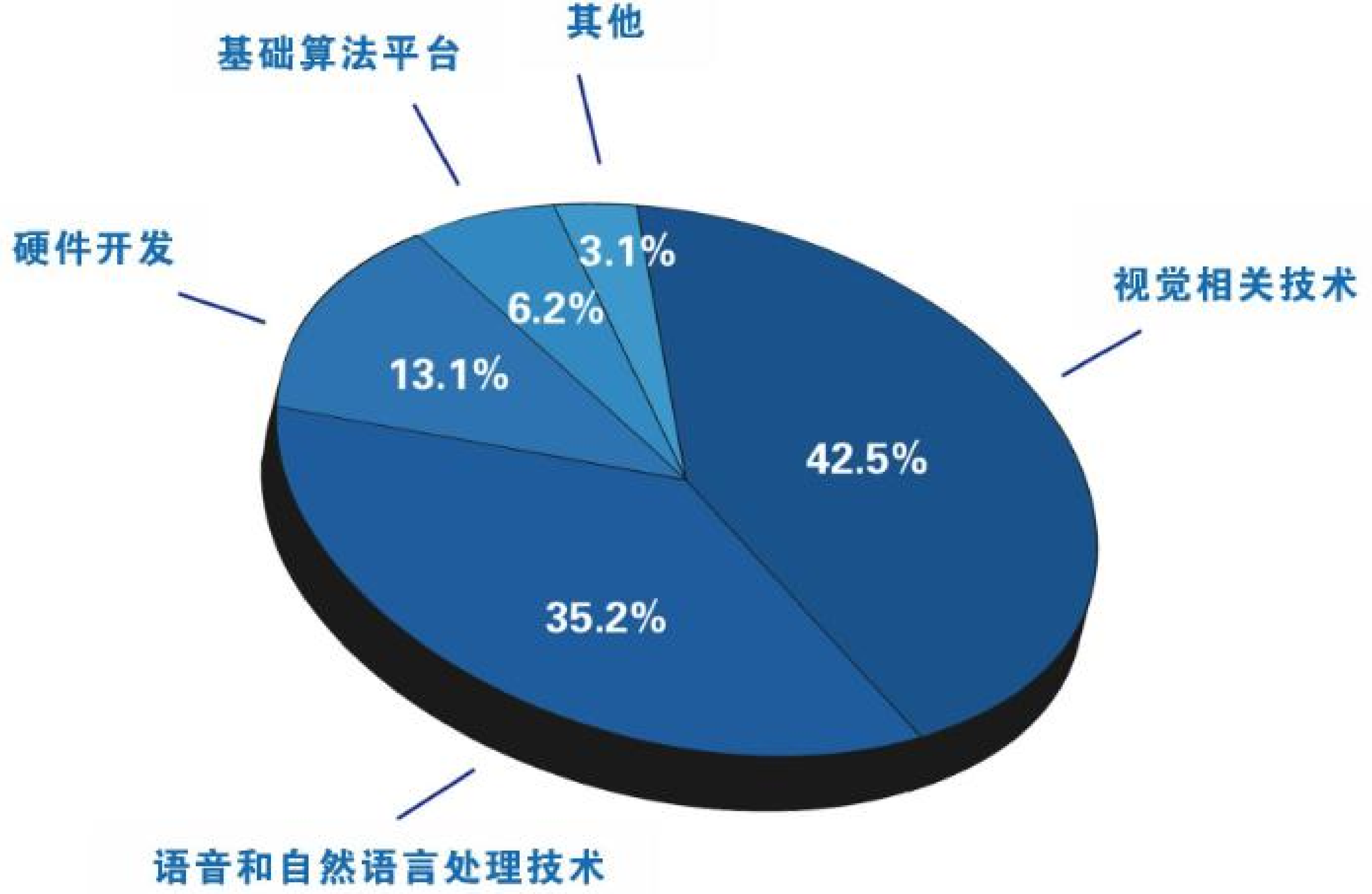


下一页 上一页


智能社会
智能升级
数字政府
智慧城市
智慧医疗
智慧教育
智慧办公
智慧交通
智慧工业
人工智能新基建
开放平台
电子政务
WeCity
健康码
远程教育
办公协同
乘车码
工业质检
服务平台
语音技术
图像技术
视频技术
AR/VR
知识图谱
NLP技术
评测技术
算法平台
算法核心框架
深度学习工具组件
AI 中台能力
云基支撑
弹性计算
无限存储
加速网络
全域安全

下一页 上一页

国内人工智能技术分布



页

人工智能优缺点

优势

1 能处理分析大规模复杂数据，
2 计算能力强大
3 应用范围广泛。

缺点

1 计算、存储代价大
2 缺乏理论保证，属于黑箱模型
3 依赖大量的训练样本

下一页 上一页

1.3 优化问题定义

$$\begin{aligned} &\underset{x}{\text{minimize}} && f(x) \\ &\text{subject to} && h_i(x)=0, i=1,2,\cdots,m \\ & && g_j(x)\geq 0, j=1,2,\cdots,n \end{aligned}$$

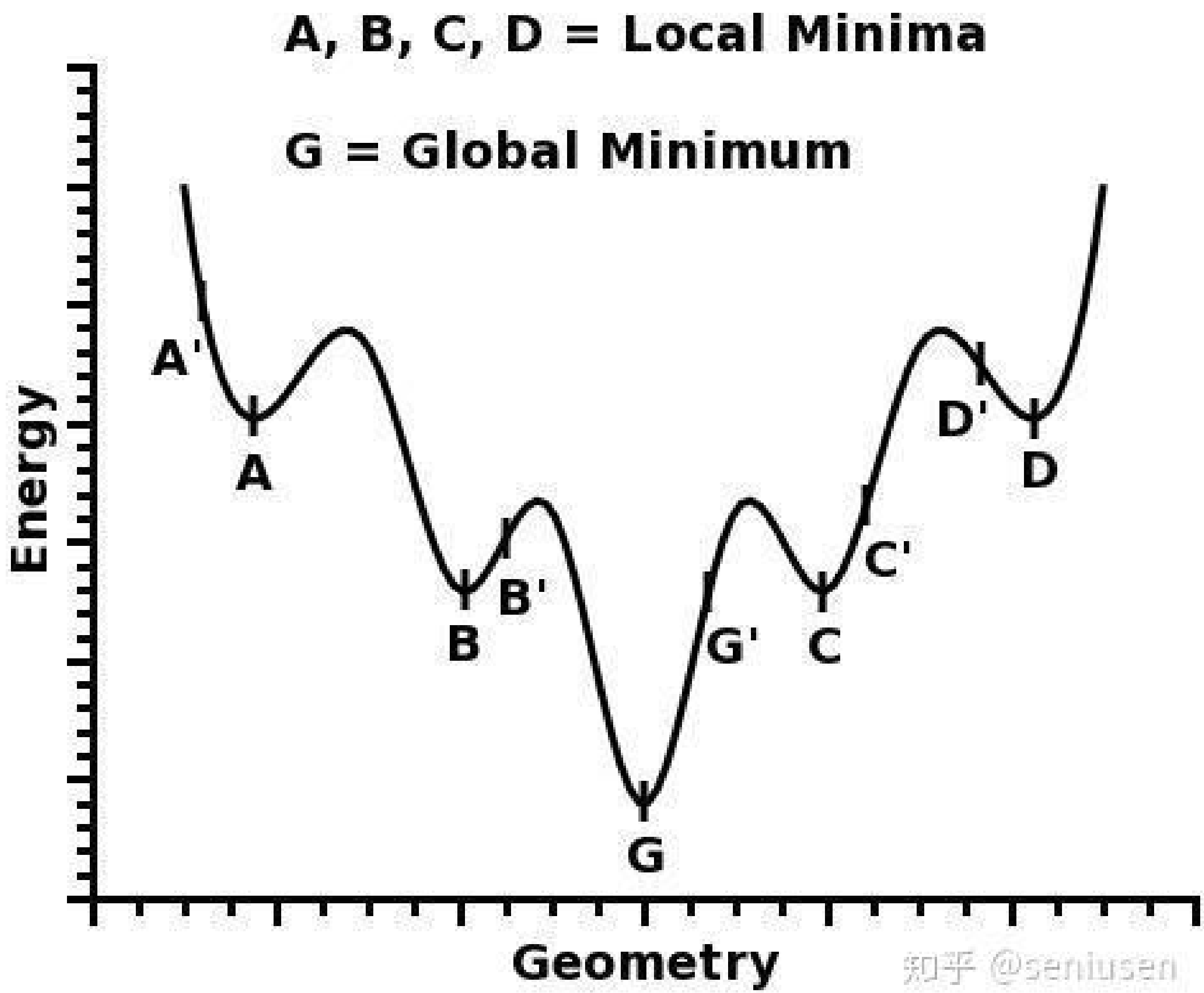
等式约束

不等式约束

目标函数f(x)，优化未知量x
minimize/maximize：min/max
subject to : s.t.

或者等价于 $x^* = \arg\min_x f(x)$

最优解与局部最优解


A, B, C, D = Local Minima
G = Global Minimum
Energy
Geometry
A'
A
B
B'
G
G'
C
C'
D'
D
知乎 @seniusen

$$
\begin{aligned}
&\underset{x}{\text{minimize}} && f(x) \\
&\text{subject to} && h_i(x)=0, i=1,2,\cdots,m \\
& && g_j(x)\geq 0, j=1,2,\cdots,n
\end{aligned}
$$

定义 (整体最优解) 若$x^*\in D$,对于一切$x\in D$恒有$f(x^*)\leq f(x)$,则称x^*为最优化问题(P)的**整体最优解**.

若$x^*\in D, x\neq x^*$, 恒有$f(x^*) < f(x)$,则称x^*为最优化问题(P)的**严格整体最优解**.

定义(局部最优解) 若$x^*\in D$,存在x^*的某邻域$N_\varepsilon(x^*)$,使得对于一切$x\in D\cap N_\varepsilon(x^*)$,恒有$f(x^*)\le f(x)$,则称为最优化问题(P)的局部最优解,其中$N_\varepsilon(x^*)=\{x|\ \|x-x^*\|<\varepsilon,\varepsilon>0\}$.

- 显然,整体最优解一定是局部最优解,而局部最优解不一定是整体最优解.
- x^*对应的目标函数值$f(x^*)$称为最优值，记为f^*.

下一页 上一页

1.4 优化问题分类：按约束条件分类

1. 无约束优化问题：

$$\boxed{\min_{x} f(x)}$$

2. 等式约束优化问题：

$$\min_{\mathbf{x}\in\mathbb{R}^2} f(\mathbf{x}) \quad \text{subject to} \quad h(\mathbf{x})=0$$

3. 不等式约束优化问题：

$$\min_{\mathbf{x}\in\mathbb{R}^2} f(\mathbf{x}) \quad \text{subject to} \quad g(\mathbf{x})\le 0$$

4.混合约束优化问题：

$$\begin{aligned}
&\min_{x} f(x)\\
&\text{s.t.}\quad h_i(x)=0, i=1,2,\cdots,m\\
&\qquad\ \ g_j(x)\ge 0, j=1,2,\cdots,n
\end{aligned}$$

按解分类

1. 凸优化问题：

局部最优解==全局最优解

传统凸优化算法进行求解

2. 非凸优化问题：

含有若干个局部最优解

智能算法进行求解

1.5 传统优化算法

- 梯度下降法
- 牛顿迭代法
- 算子分裂法
- 对偶算法
- 对数障碍法

梯度下降算法

问题：求函数f(x)的最小值

$$\min_{x} f(x)$$

输入：初始点 x_0，固定步长 α

k = 1;

Repeat

$$x_k = x_{k-1} - \alpha \nabla f(x_{k-1})$$

$$k = k + 1$$

Until (满足终止条件)

传统优化算法优缺点

- 有较为明确的问题和条件描述，
- 很多属于凸优化范畴，有唯一明确的全局最优点；对于单极值问题，传统算法优于智能算法
- 一般是确定性算法，有固定的结构和参数，计算复杂度和收敛性；
- **理论较为完善，计算量小。**
- **收敛速度快。**
- **具有确定的终止准则。**

缺点：

1. **仅能求出优化问题的局部最优解。**
2. **求解的结果强烈依赖于初始值。**
3. **应用范围有很大局限性**

1.6. 智能算法

受人类智能、生物群体或自然现象启发，研究设计算法

主要包括：

① 模仿生物进化机制的遗传算法

② 模拟生物免疫系统的免疫算法

③ 模拟蚂蚁寻径行为的蚁群算法

④ 模拟鸟群群体行为的粒子群算法

⑤ 模拟固体物质退火过程的模拟退火算法

⑥ 模拟动物神经网络行为的神经网络算法

1.6.1 遗传算法

Genetic Algorithm, GA

模拟生物在自然环境中的遗传和进化过程而形成的全局优化搜索算法。

生物遗传概念	遗传算法中的作用
适者生存	在算法停止时，最优目标值的解有最大的可能被保留
个体(individual)	解
染色体(chromosome)	解的编码（字符串，向量等）
基因(gene)	解中每一分量的特征（如各分量的值）
适应性(fitness)	适应函数值
群体(population)	选定的一组解（其中解的个数为群体的规模）
种群(reproduction)	根据适应函数值选取的一组解
交配(crossover)	通过交配原则产生一组新解的过程
变异(mutation)	编码的某一个分量发生变化的过程

下一页 上一页

1.6.2 免疫算法

于1973年提出, Immune Algorithm, IA

模仿生物免疫机制,

生物遗传概念	遗传算法中的作用
适者生存	在算法停止时，最优目标值的解有最大的可能被保留
个体(individual)	解
染色体(chromosome)	解的编码（字符串，向量等）
基因(gene)	解中每一分量的特征（如各分量的值）
适应性(fitness)	适应函数值
群体(population)	选定的一组解（其中解的个数为群体的规模）
种群(reproduction)	根据适应函数值选取的一组解
交配(crossover)	通过交配原则产生一组新解的过程
变异(mutation)	编码的某一个分量发生变化的过程

下一页 上一页

1.6.3 蚁群算法

Ant Colony Optimization, ACO

蚁群算法的操作：

模仿蚂蚁在没有任何提示下找到：从巢穴到食物源的最短路径， 建立的一种基于种群的启发式随机搜索算法。

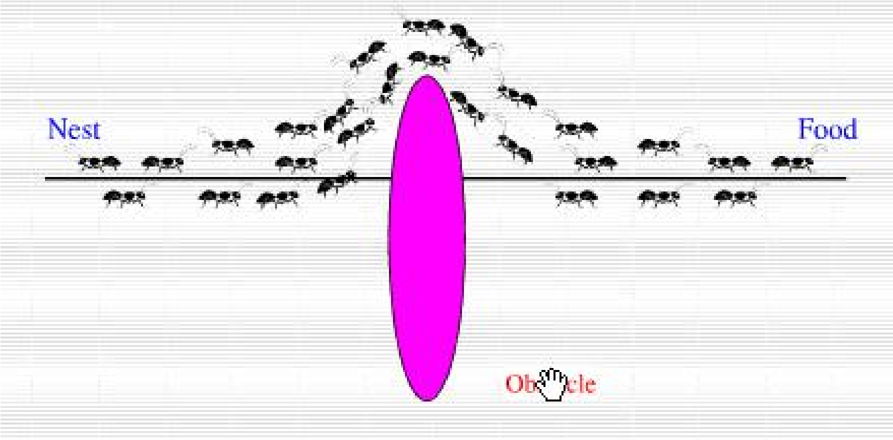


图4 蚂蚁最终绕过障碍物找到最优路径

1.6.4 粒子群算法

Particle Swarm Optimization, PSO

粒子群算法的操作：

模拟鸟群觅食过程中迁徙和群聚行为，

通过群体中个体之间的协作和信息共享来寻找最优解

$$v_{id}^{k+1} = wv_{id}^{k} + c_1 rand()(p_{id} - x_{id}^{k}) + c_2 rand()(gbest_d - x_{id}^{k})$$

$$x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1} \qquad i = 1,2,\cdots,m;\ d = 1,2,\cdots,D$$

其中，w称为惯性权重，c_1和c_2为两个正常系数，称为加速因子。

$$v_{id}^{k+1} = w v_{id}^{k} + c_1 rand()(p_{id} - x_{id}^{k}) + c_2 rand()(g_d - x_{id}^{k})$$

$$x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1} \qquad i = 1,2,\cdots,m;\ d = 1,2,\cdots,D$$

“惯性部分”，对自身运动状态的信任

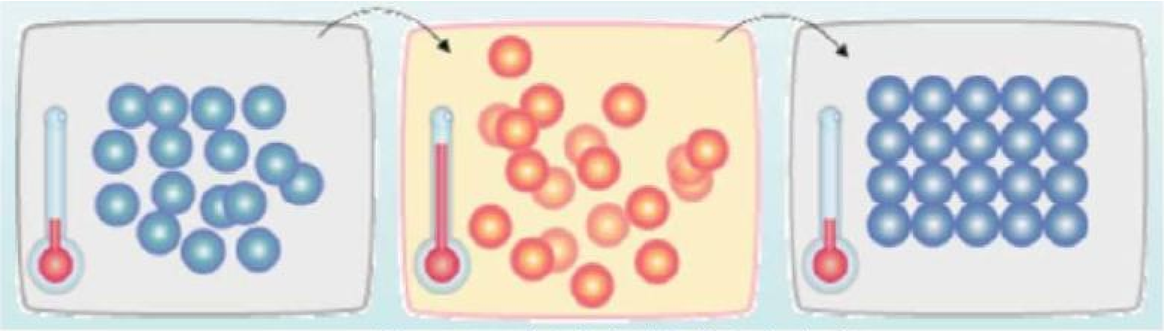
“认知部分”，对粒子本身的思考，即来源于自己经验的部分

“社会部分”，粒间子的信息共享，来源于群体中的其它优秀微粒的经验

1.6.5 模拟退火算法

Simulated Annealing, SA

模拟退火算法的操作：



物理退火过程的发展阶段

1. 加温过程。其目的是增强粒子的热运动，使其偏离平衡位置。当温度足够高时，固体将溶解为液体，溶解过程与系统的熵增过程联系，系统能量也随温度的升高而增大，使得每一粒子的状态都具有充分的随机性。

2. 等温过程。物理学的知识告诉我们，对于与周围环境交换热量而温度不变的封闭系统，系统状态的自发变化总是朝自由能减少的方向进行，当自由能达到最小时，系统达到平衡态。

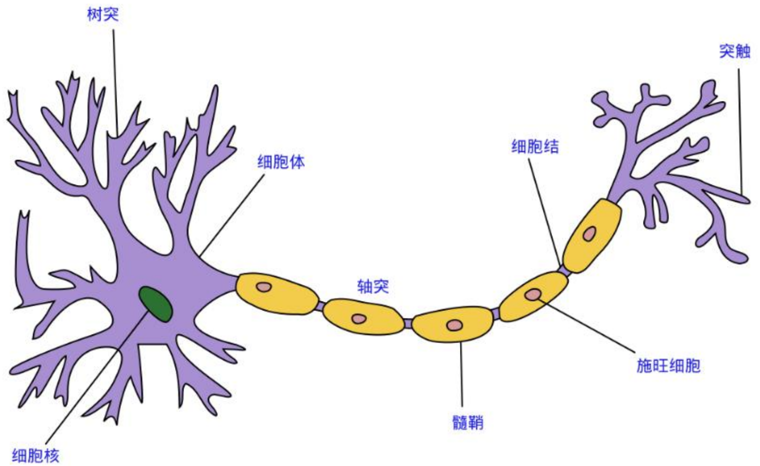
3. 冷却过程。目的是使粒子的热运动减弱并渐趋有序，系统能量逐渐下降，从而得到低能的晶体结构。在常温时达到基态，内能减为最小。

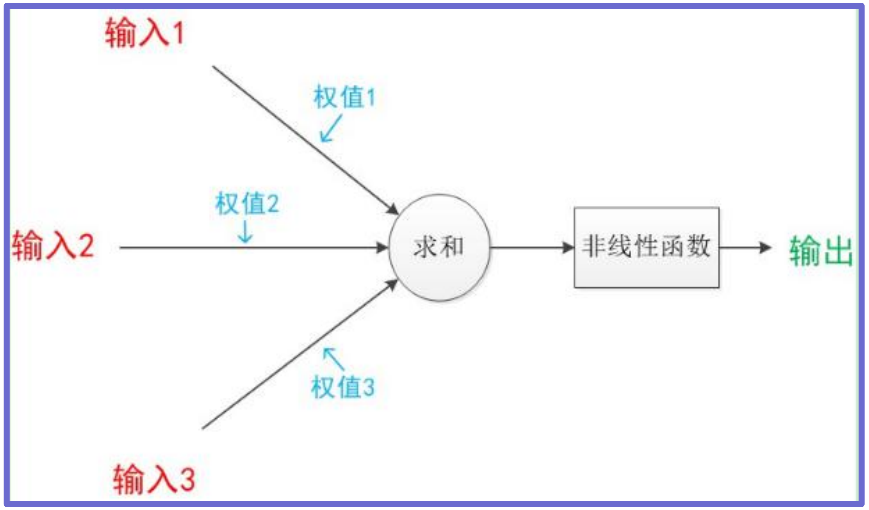
上一页

1.6.6 人工神经网络算法

Artificial Neural Network，ANN

模拟动物与人脑的网络结构， 建立的一种信息处理模型。



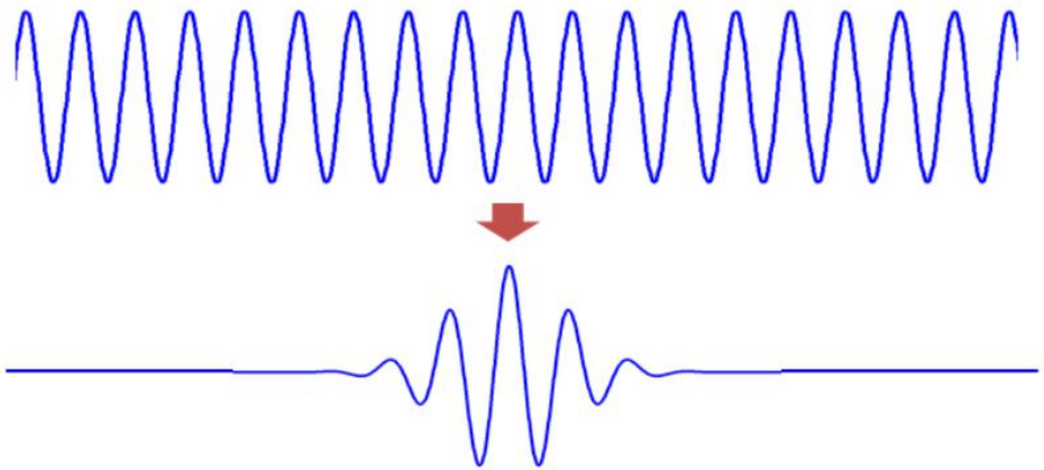


1.6.7 小波分析

小波做的改变就在于，将无限长的三角函数基换成了有限长的会衰减的小波基。

小波变换

$$F(w)=\int_{-\infty}^{\infty} f(t)*e^{-iwt}dt \Rightarrow WT(a,\tau)=\frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} f(t)*\psi(\frac{t-\tau}{a})dt$$



尺度a控制小波函数的**伸缩**，**平移量** τ控制小波函数的**平移**。

智能算法

- 针对多极值问题，通过有效设计可以实现找到全局最优点，有很大应用空间和改进可能。
- 大多属于随机启发性算法，能定性分析却难定量证明，收敛速度慢，计算复杂度较高。
- “没有办法的办法”
- 适合于求解复杂的优化问题。

缺点

1. 收敛速度慢。
2. 局部搜索能力差。
3. 控制变量较多。
4. 无确定的终止准则。

下一页 上一页

第2章　数学基础知识

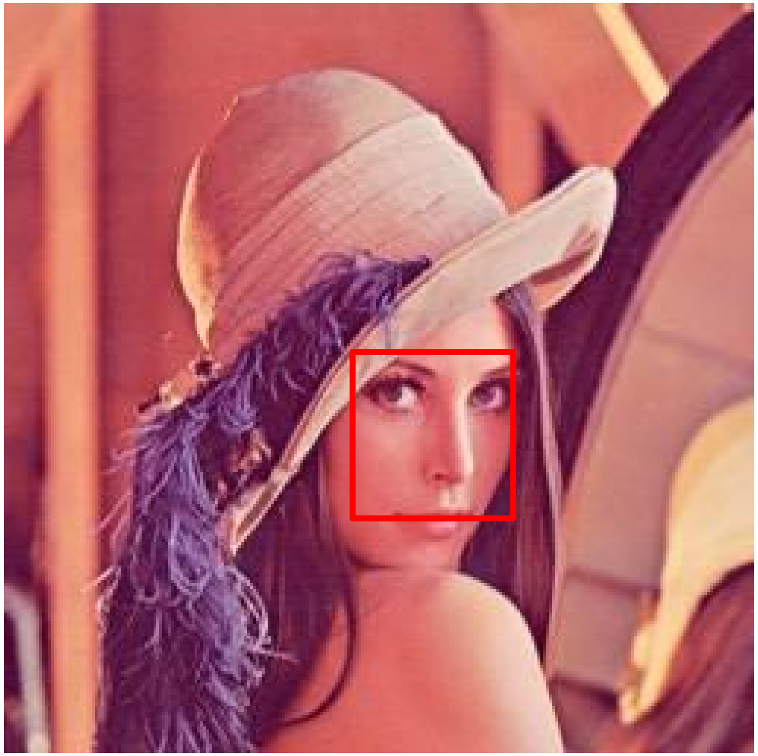
2.1 数字图像数据

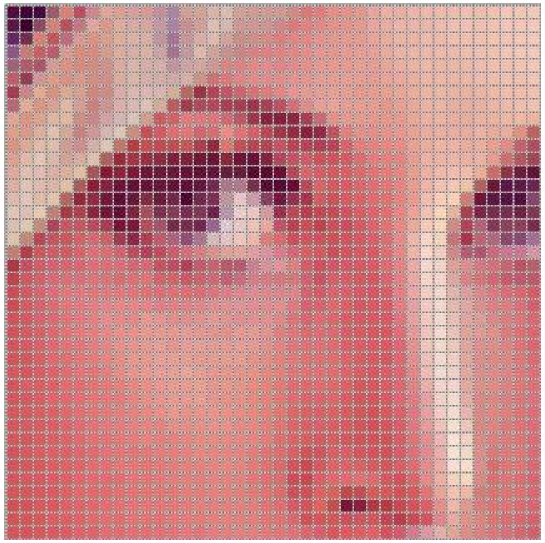
2.2 导数，梯度，散度

2.3 范数

2.1 数字图像数据

- 数字图片，通过像机等拍摄获取，
- 数据类型： jpg, png, gif....
- 规则的格子构成的二维数据 x,y





2.1.1 数字图像的表示

$$
f(x,y)=\begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix}
$$

这个表达式的右侧定义 了一幅数字图像。矩阵 中的每个元素称为图像像素。

1. 大小为 M*N的图像：二维函数 f(x, y),
2. (x, y)为坐标，f(x, y)为坐标值，也即像素值

3. 灰度图：M*N, f(i, j) = c
   彩色图：M*N*3, f(i, j)＝（ r ， g ， b ）

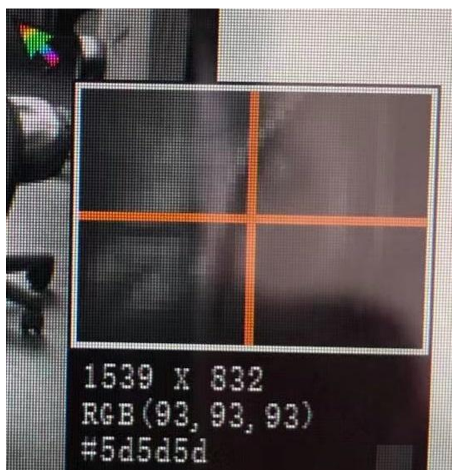
4. 其中，像素值的大小范围为[0, 255]的整数值

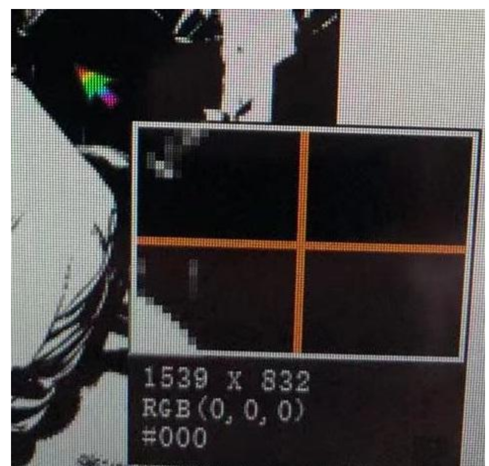
彩色图、灰度图、二值图

[r,g,b]--[0,255]

[c]--[0,255]

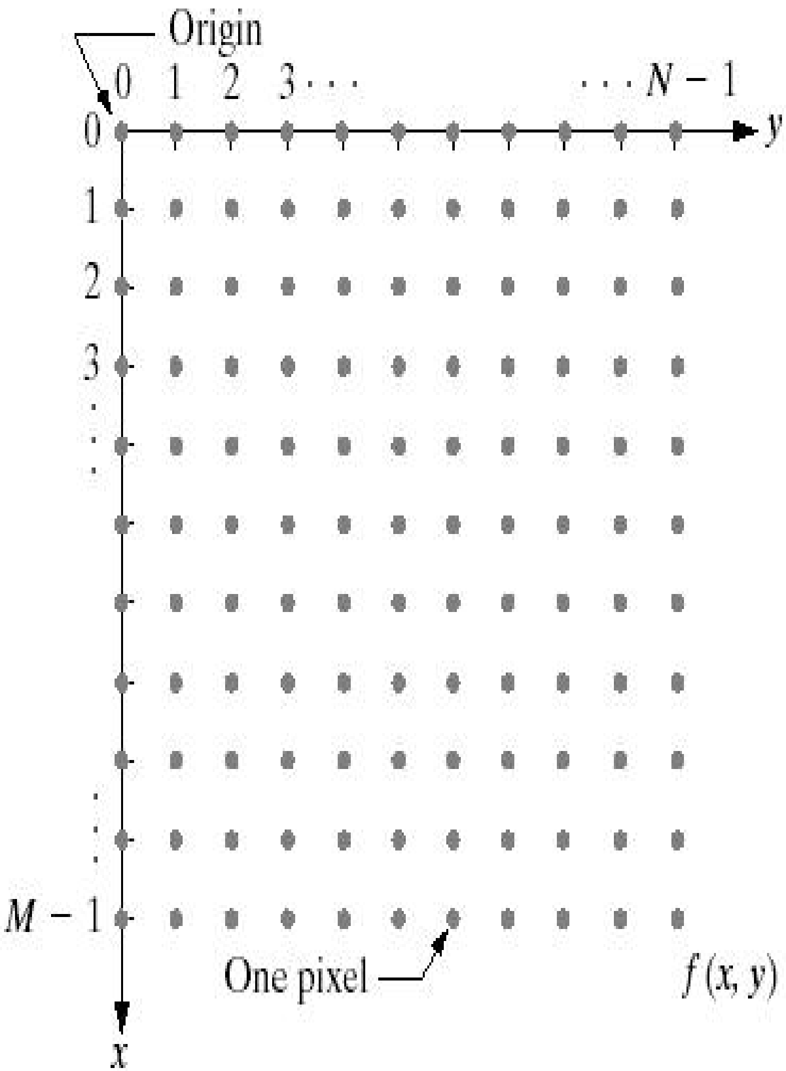
0,1





下一页 上一页

2.1.2 像素坐标系：



左图为，M*N图像的坐标系

1、坐标原点位于左上角

2、数据先沿x轴增加

3、然后再沿y轴增加

4、坐标轴为整数

5、matlab中原点为(1,1)

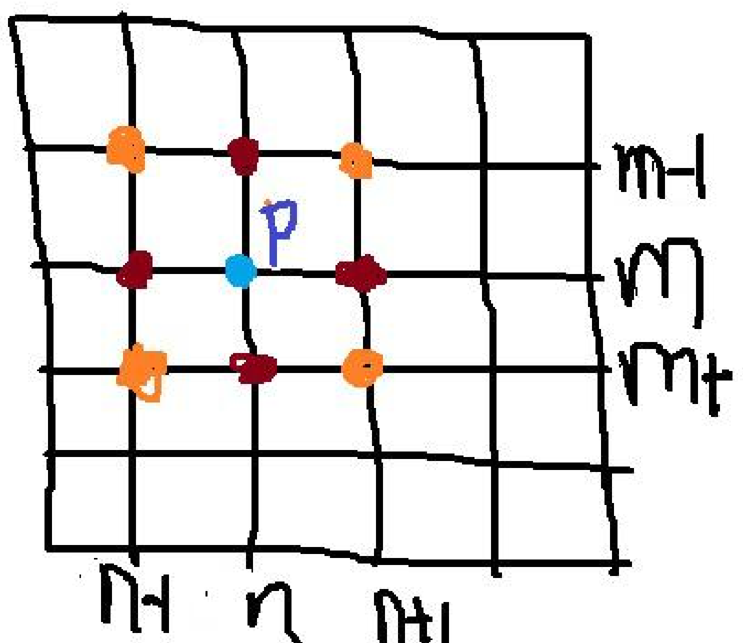
2.1.3 像素间的关系与距离

- 像素p(m,n)的相邻像素

4邻域 $N_4(p)$：(m+1, n), (m-1, n), (m, n+1), (m, n-1)

对角邻域 $N_D(p)$：(m+1, n+1), (m+1, n-1), (m-1, n+1), (m-1, n-1)

8邻域 $N_8(p)$：$N_4(p)$ + $N_D(p)$



下一页 上一页

像素间距离的度量

像素P(x, y), Q(s, t)间的距离

- 欧氏距离：

$$D_e(p,q)=\left[(x-s)^2+(y-t)^2\right]^{\frac{1}{2}}$$

- D_4距离（城市街区距离）：

$$D_4(p,q)=|x-s|+|y-t|$$

- D_8距离（棋盘距离）：

$$D_8(p,q)=\max\left(|x-s|,|y-t|\right)$$

2.1.4 数字图像的 matlab命令

- 读取图像：I = imread('scene.jpg');
- 显示图像：imshow(I);
- 获取图像的大小：[m,n,r] = size(I);
- 彩图转灰度图：gray_I = rgb2gray(I);
- 图像数据类型转换：double(I);
- RGB转HSV：I_h=rgb2hsv(I);
- 彩色图转二值图像：

  thresh_I = graythresh(I);

  im2_I = im2bw(I, thresh_I);

2.2 图像数据的微分算子

标量定义：只有大小，没有方向，常数

向量定义：有大小，也有方向

n维向量写成行的形式，称为行向量，记为

$$\alpha^T = (a_1, a_2, \cdots, a_n),$$

写成列的形式，称为列向量，记为

$$\alpha = \begin{pmatrix} a_1 \\ a_2 \\ \cdots \\ a_n \end{pmatrix}.$$

下一页　上一页

2.2.1 向量内积

向量内积：常数

$$\vec{i}=(1,0,0) \qquad \vec{j}=(0,1,0) \qquad \vec{k}=(0,0,1)$$

$$\boldsymbol{a}=(x_1,y_1,z_1)=x_1\vec{i}+y_1\vec{j}+z_1\vec{k},$$

$$\boldsymbol{b}=(x_2,y_2,z_2)=x_2\vec{i}+y_2\vec{j}+z_2\vec{k}$$

$$<\mathrm{a},\mathrm{b}>=\mathrm{a}\cdot\boldsymbol{b}=\boldsymbol{x}_1\boldsymbol{y}_1+\boldsymbol{x}_2\boldsymbol{y}_2+\boldsymbol{x}_3\boldsymbol{y}_3$$

对任意的 $x,y\in\mathbf{R}^n$，

$$\langle x,y\rangle=x^Ty=\sum_{i=1}^{n}x_iy_i.$$

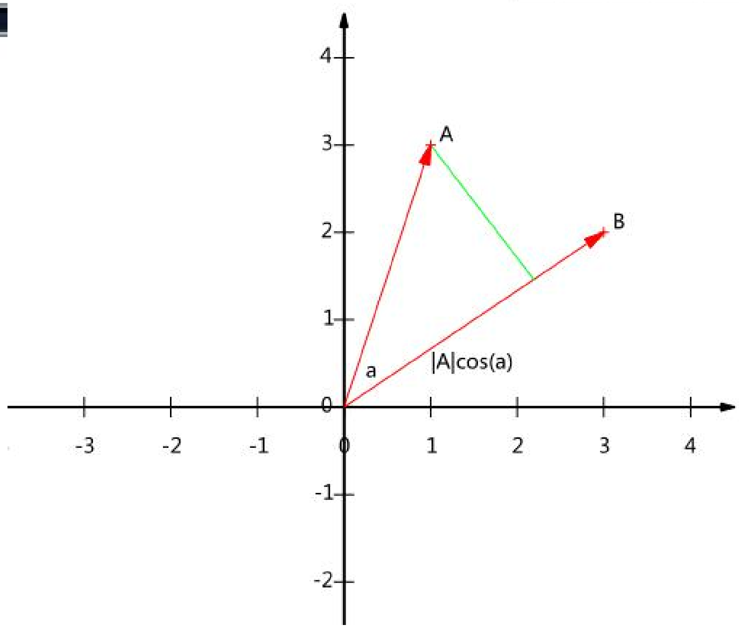
投影

- 内积：二维向量A,B内积定义

$$A \cdot B = |A||B|\cos\theta$$

- 投影：假设B为单位向量，|B|=1，

$$A \cdot B = |A|\cos\theta$$



- A与B的**内积值**等于A向B所在直线投影的矢量长度

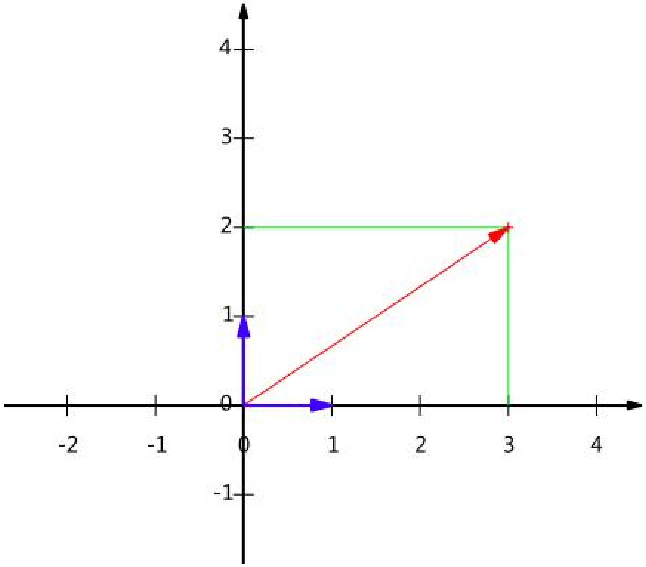
坐标系与向量

- 右图中二维坐标系的基： $\vec{e}_1=(1,0),\vec{e}_2=(0,1)$
- 向量(3,2)完整表示为：

$$(3,2)=3\vec{e}_1+2\vec{e}_2$$

- 在x轴方向的投影为3，在y轴方向的投影为2
- 上式又可写为

$$(3,2)=<(3,2),\vec{e}_1>\vec{e}_1+<(3,2),\vec{e}_2>\vec{e}_2$$



因此，对于n维向量 $\vec{a}=(a_1,a_2,\cdots,a_n)$,

有 $\vec{a}=\sum_{i=1}^{n}<\vec{a},e_i>e_i$

2.2.2 矩阵

定义1 由 $m\times n$ 个数 $a_{ij}(i=1, 2, \cdots, m;\ j=1, 2, \cdots, n)$排成一个 m 行 n 列的矩形表称为一个 $m\times n$ 矩阵，记作

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

其中 a_{ij} 称为矩阵的第 i 行第 j 列的元素.

一般情况下，我们用大写字母 A，B，C 等表示矩阵.

$m\times n$矩阵A简记为 $A=(a_{ij})_{m\times n}$ 或记作 $A_{m\times n}$.

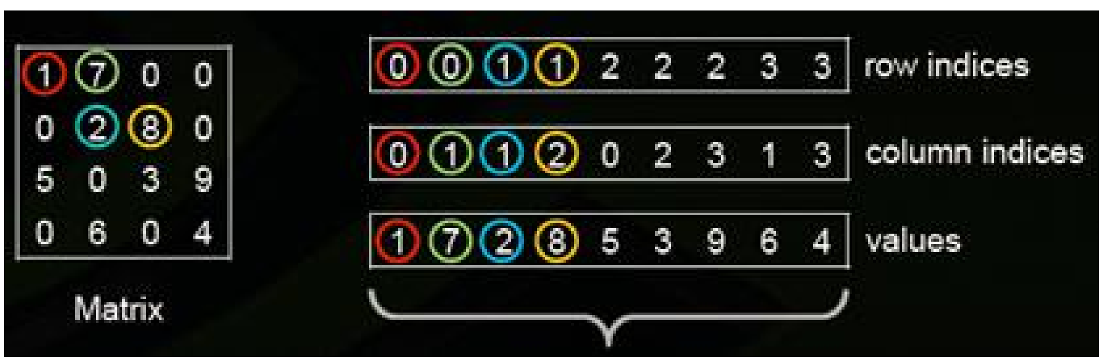
稀疏矩阵(Sparse Matrix)

定义：矩阵中非零元素的个数远远小于矩阵元素的总数

$$M=\begin{bmatrix} 2 & 0 & 0 & 0 & 6 & 0 & 0 & 7 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 \\ 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

稀疏矩阵示意图

存储方式：用三元组来表示，分别是（行号，列号，数值）




```
a = [1 1 1 2 3 3 4 5 6];
b = [1 5 8 3 3 7 6 4 2];
v = [2 6 7 1 2 3 8 5 9];
A = sparse(a,b,v);
```

矩阵内积

1. 矩阵 A, B 内积定义：对应元素相乘求和

$$\mathrm{A}\bullet\mathrm{B}=\sum_{ij}a_{ij}b_{ij}$$

2. 矩阵A的迹(trace)：A的主对角线（从左上方至右下方的对角线）上各个元素的总和：

$$\mathrm{tr}A=\sum_{i=1}^{n}A_{ii}.$$

矩阵分解

- 对称特征值分解，矩阵为方阵

假设 $A \in \mathbf{S}^n$，即 A 是实对称 $n \times n$ 矩阵。那么 A 可以因式分解为

$$A = Q \Lambda Q^T,$$

其中 $Q \in \mathbf{R}^{n \times n}$ 是正交矩阵，即满足 $Q^T Q = I$，

而 $\Lambda = \mathbf{diag}(\lambda_1, \cdots, \lambda_n)$ ：λ_i 是 A 的特征值

矩阵 $A \in \mathbf{R}^{n \times n}$ 被称为正交矩阵的条件是

$$A^T A = I, \text{即 } A^{-1} = A^T$$

- 奇异值分解：矩阵为任意阶矩阵

假设 $A \in \mathbf{R}^{m\times n}$，$\mathbf{rank}\, A = r$。那么 A 可以因式分解为

$$A = U\Sigma V^T,$$

其中 $U \in \mathbf{R}^{m\times r}$ 满足 $U^TU = I$， $V \in \mathbf{R}^{n\times r}$ 满足 $V^TV = I$，
$\Sigma = \mathbf{diag}(\sigma_1, \cdots, \sigma_r)$ σ_i 则称为奇异值。

求$\boldsymbol{U},\boldsymbol{V},\boldsymbol{\Sigma}$分别是什么？

如何求出分解后的三个矩阵？ U,V,Σ

- 构造对称方阵： $\boxed{A^T A}$
- 对其进行特征分解：

$$\boxed{A^T A v_i = \lambda_i v_i \quad \text{记}\ V' = [v_1 \cdots v_n]}$$

- 同样构造对称方阵： $\boxed{A\ A^T}$
- 对其进行特征分解：

$$\boxed{A A^T u_i = \lambda_i u_i \quad \text{记}\ U' = [u_1 \cdots u_m]}$$

注：
矩阵与其转置矩阵的特征值相同

如何求出分解后的三个矩阵? U,V,Σ

- **需要证明A分解的U,V为以上定义的U,V**

$$\because\ A = U\Sigma V^T,$$
$$\therefore\ A^T = V\Sigma^T U^T \Rightarrow A^T A = V\Sigma^2 V^T$$

$\therefore\ A^T A$的特征向量组V即是SVD中的V矩阵

同理AA^T的特征向量组U即是SVD中的U矩阵

Σ: $\sigma_{\mathrm{i}} = \sqrt{\lambda_i}$

注:

矩阵与其转置矩阵的特征值相同

奇异值分解可以写成

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T,$$

其中 $u_i \in \mathbf{R}^m$ 是左奇异向量，$v_i \in \mathbf{R}^n$ 是右奇异向量。

例： 求矩阵的SVD分解：

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}$$

首先求出$A^T A$和AA^T

$$\mathbf{A}^\mathbf{T}\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

$$\mathbf{A}\mathbf{A}^\mathbf{T} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

进而求出A^TA的特征值和特征向量：

$$\lambda_1 = 3; v_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}; \lambda_2 = 1; v_2 = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

接着求AA^T的特征值和特征向量：

$$\lambda_1 = 3; u_1 = \begin{pmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix}; \lambda_2 = 1; u_2 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{pmatrix}; \lambda_3 = 0; u_3 = \begin{pmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}$$

用$\sigma_i = \sqrt{\lambda_i}$直接求出奇异值为$\sqrt{3}$和1.

$$A = U\Sigma V^T = \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

SVD的性质：矩阵的秩为k时

- 矩阵的奇异值个数很少，可如下逼近：

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}V^T_{n\times n} \approx U_{m\times k}\Sigma_{k\times k}V^T_{k\times n}$$