几何体的向量表示

已知二维平面上两定点 A(5, 1), B(2, 3), 给出线段 AB 的方程如下:

$$\begin{cases} x_1 = \theta * 5 + (1-\theta) * 2 \\ x_2 = \theta * 1 + (1-\theta) * 3 \end{cases}$$

如果将点 A 堪称向量 a,点 B 看成向量 b,则线段 AB 的向量表示为:

$$\vec{x} = \theta\vec{a} + (1-\theta) * \vec{b}, \qquad \theta \in [0,1]$$

直线的向量表示是

$$\vec{x} = \theta\vec{a} + (1-\theta) * \vec{b}, \qquad \theta \in R$$

故我们可以将类似的想法推广到多维平面

三角形:

$$\vec{x} = \theta_1\overrightarrow{a_1} + \theta_2\overrightarrow{a_2} + \theta_3\overrightarrow{a_3}, \qquad \theta_i \in [0,1] \text{ and } \sum \theta_i = 1$$

二维平面:

$$\vec{x} = \theta_1\overrightarrow{a_1} + \theta_2\overrightarrow{a_2} + \theta_3\overrightarrow{a_3}, \qquad \theta_i \in R \text{ and } \sum \theta_i = 1$$

超几何体的向量表示:

$$\vec{x} = \theta_1\overrightarrow{a_1} + \theta_2\overrightarrow{a_2} + \cdots + \theta_k\overrightarrow{a_k}, \qquad \theta_i \in [0,1] \text{ and } \sum \theta_i = 1$$

超平面的向量表示:

$$\vec{x} = \theta_1\overrightarrow{a_1} + \theta_2\overrightarrow{a_2} + \cdots + \theta_k\overrightarrow{a_k}, \qquad \theta_i \in R \text{ and } \sum \theta_i = 1$$

凸集

集合 C 内任意两点间的线段也均在集合 C 内,则称集合 C 为凸集,数学定义为:

$$\forall x_1, x_1, \dots x_k \in C, \theta_i \in [0,1]\ and \sum \theta_i = 1, then\ x = \sum_{i=1}^{k} \theta_i x_i \in C$$

上面凸集的定义中使用到了超几何体的向量表示,<u>且凸集的交集仍是凸集</u>.

凸函数

凸函数定义为:

$$f: C \subseteq R^n -> R^1, C.x_1, x_2 \in C,:$$

$$f(\alpha_1 x_1 + \alpha_2 x_2) <= \alpha_1 f(x_1) + \alpha_2 f(x_2) \quad \sum a_i = 1, a_i >= 0$$

则称 f(x)为定义在凸集 C 上的凸函数.

严格凸函数定义为:

$$f: C \subseteq R^n -> R^1, C.x_1, x_2 \in C,:$$

$$f(\alpha_1 x_1 + \alpha_2 x_2) < \alpha_1 f(x_1) + \alpha_2 f(x_2) \quad \sum a_i = 1, a_i >= 0$$

则称 f(x)为定义在凸集 C 上的严格凸函数.

根据一阶导数判断函数凸性:

设 f(x)为定义在凸集 R 上,且为具有连续一阶导数的函数,则 f(x)在 R 上为凸函数的充要条件是对凸集 R 内任意不同两点,不等式

$$f(x_2) >= f(x_1) + (x_2 - x_1)^T \nabla f(x_1)$$

恒成立.

根据二阶导数判断函数凸性:

设 f(x)为定义在凸集 R 上,且为具有连续二阶导数的函数,则 f(x)在 R 上为凸函数的充要条件是 Hesse 矩阵在 R 上处处半正定.

几何含义:

函数任意A_1, A_2之间的部分位于A_1, A_2连线的下方.

凸函数优点

凸函数的局部极小值是全局极小值.在机器学习中,我们只需要将非凸函数转化为凸函数,便可直接通过梯度下降或者牛顿法得出问题的全局极值.

詹森不等式

f(x)为凸函数,其中 E(x)是 x 的期望,我们有

$$f(E(x)) <= E(f(x))$$

无约束优化

解决方法就是对变量求导.

等式约束优化

1. 可以使用消元法,因为是登时,可以代入原式
2. 拉格朗日乘子法