

# AI2613: 随机过程

上海交通大学

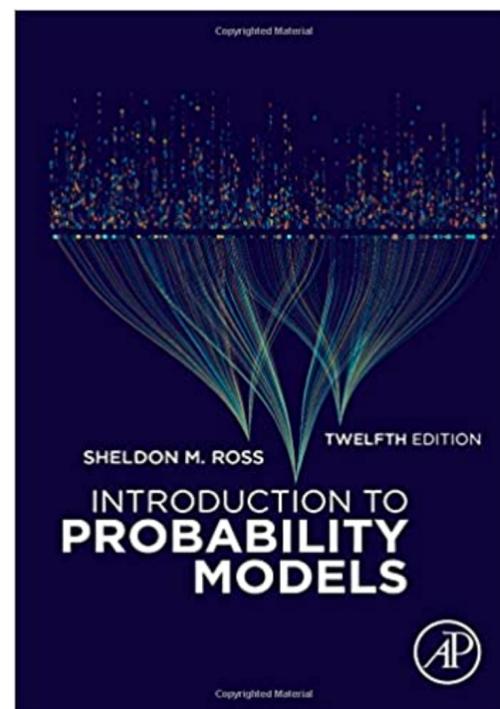
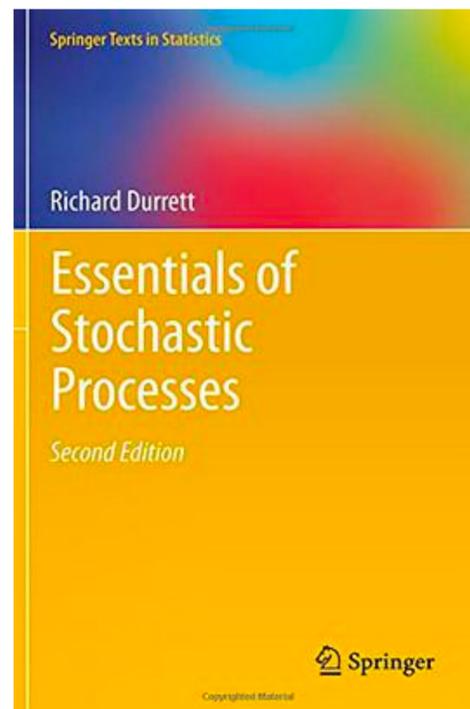
张驰豪

2021年2月22日

# 课程信息

- 每周一 12:55 - 15:40 @ 东上院 107
- 任课教师：张驰豪 (<http://chihaozhang.com>, [chihao@sjtu.edu.cn](mailto:chihao@sjtu.edu.cn))
- 助教：邱国良 ([guoliang.qiu@sjtu.edu.cn](mailto:guoliang.qiu@sjtu.edu.cn))  
陈厚双 ([chenhoushuang@sjtu.edu.cn](mailto:chenhoushuang@sjtu.edu.cn))
- 课程主页：<http://chihaozhang.com/teaching/SP2021spring>
- Office hour: 每周一晚上7点-10点@软件学院1402-2

- 每两周一次作业
- 最终成绩 = 70% \* 平时作业 + 30% \* 期末考试
- 参考资料：手写讲义 + pdf讲义 + 课程视频@canvas



CONTENTS Page 1

### Contents

<b>1 Markov chains</b>	<b>5</b>
1.1 Specifying and simulating a Markov chain	5
1.2 The Markov property	8
1.3 "It's all just matrix theory"	9
1.4 The basic limit theorem of Markov chains	10
1.5 Stationary distributions	12
1.6 Irreducibility, periodicity, and recurrence	15
1.7 An aside on coupling	25
1.8 Proof of the Basic Limit Theorem	27
1.9 A SLLN for Markov chains	31
1.10 Exercises	36
<b>2 Markov Chains: Examples and Applications</b>	<b>43</b>
2.1 Branching Processes	43
2.2 Time Reversibility	46
2.3 More on Time Reversibility: A Tandem Queue Model	49
2.4 The Metropolis method	53
2.5 Simulated annealing	57
2.5.1 Description of the method	58
2.5.2 The Main Theorem	65
2.6 Ergodicity Concepts	67
2.6.1 The Ergodic Coefficient	68
2.6.2 Sufficient Conditions for Weak and Strong Ergodicity	69
2.7 Proof of Main Theorem of Simulated Annealing	71
2.8 Card Shuffling	73
2.8.1 "Top-in-at-random" Shuffle	74
2.8.2 Threshold Phenomenon	74
2.8.3 A random time to exact stationarity	76
2.8.4 Strong Stationary Times	77
2.8.5 Proof of threshold phenomenon in shuffling	78
2.9 Exercises	81

Stochastic Processes J. Chang, February 2, 2007

# 一个知乎上的问题

知乎

首页

发现

等你来答

浙江新冠疫苗一针 200 元



提问

数学

概率论

随机过程

概率论与数理统计

**和女朋友在商场走丢了，随机乱逛和守在特定地点等候，哪个相遇的概率更高？**

1. 商场的地形比较复杂，空间也比较大。
2. 和女朋友互相生气，不存在打电话或者留下记号之类的手段。
3. 虽然生气了，但双方都想用装作偶遇的方式和解，所以不会离开商场，也不会到其他层。
4. 女朋友会继续逛，不存在静止不动的可能。
5. 急，在线等。

关注问题

写回答

邀请回答

好问题 1

26 条评论

分享

收起 ^

(<https://www.zhihu.com/question/310437394>)





刘天任

CS Theory PhD (liutianren.com)

Heng Guo、Chao Xu 等 526 人赞同了该回答

假设你和女朋友都是（完美的）无头苍蝇，且商场空间有对称性的话，两个人都随机乱逛比其中一个踟着更好。

@Mather King 用平面上的布朗运动为模型。结论是双方均随机运动可使相遇时间减半。

我来补充个离散建模，也能得到同样的结果。

用二维网格上的随机漫步（random walk）来模拟随机乱逛的人。记男友在  $t$  时刻位于坐标为  $X_t$ ，女友  $t$  时刻位于坐标  $Y_t$ 。不失一般性，设  $X_0 = (0, 0)$ 。每过单位时间，两人都随机运动到旁边的一个整点。严格地说，记  $\Delta X_t := X_t - X_{t-1}$ ， $\Delta Y_t := Y_t - Y_{t-1}$ ，则  $\Delta X_1, \Delta X_2, \dots, \Delta Y_1, \Delta Y_2, \dots$  为  $\{(1, 0), (-1, 0), (0, 1), (0, -1)\}$  上的 i.i.d. 均匀分布。



刘天任

CS Theory PhD (liutianren.com)

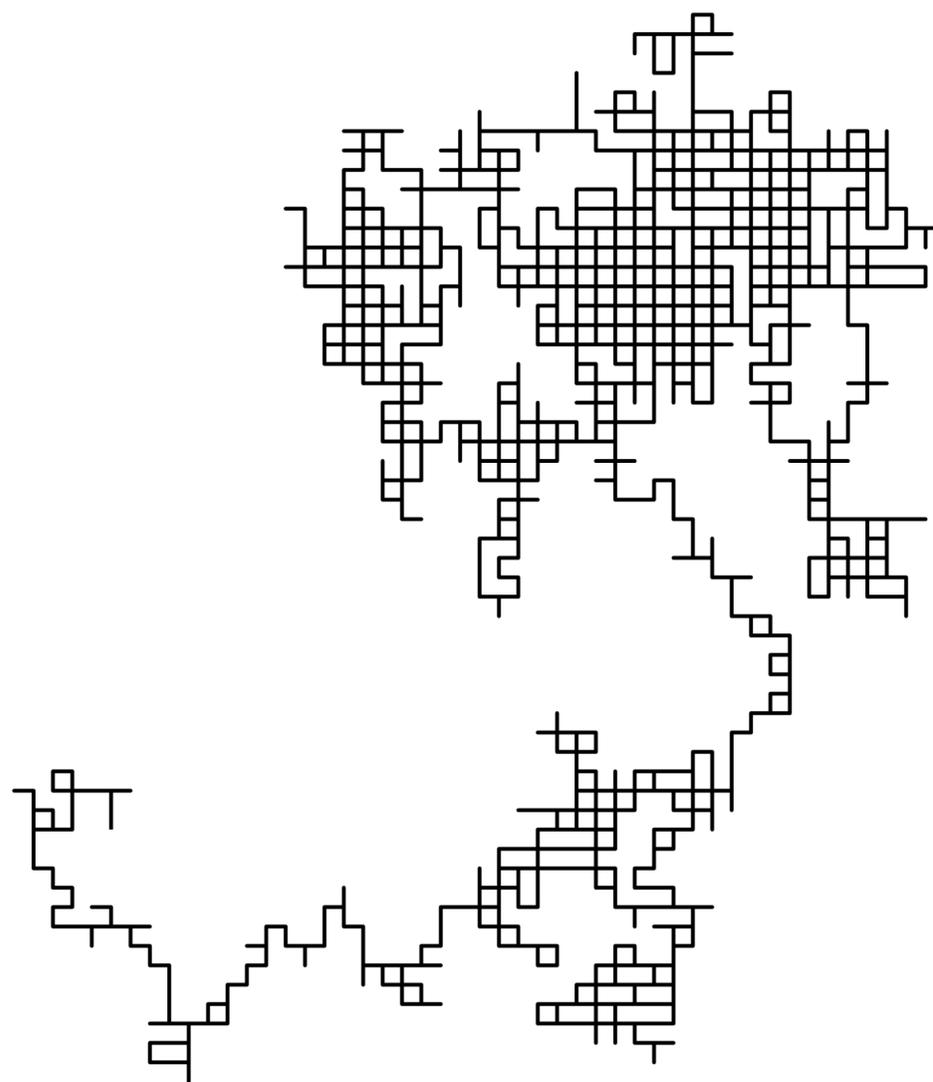
Heng Guo、Chao Xu 等 526 人赞同了该回答

假设你和女朋友都是（完美的）无头苍蝇，且商场空间有对称性的话，两个人都随机乱逛比其中一个踮着更好。

@Mather King 用平面上的布朗运动为模型。结论是双方均随机运动可使相遇时间减半。

我来补充个离散建模，也能得到同样的结果。

用二维网格上的随机漫步（random walk）来模拟随机乱逛的人。记男友在  $t$  时刻位于坐标为  $X_t$ ，女友  $t$  时刻位于坐标  $Y_t$ 。不失一般性，设  $X_0 = (0, 0)$ 。每过单位时间，两人都随机运动到旁边的一个整点。严格地说，记  $\Delta X_t := X_t - X_{t-1}$ ， $\Delta Y_t := Y_t - Y_{t-1}$ ，则  $\Delta X_1, \Delta X_2, \dots, \Delta Y_1, \Delta Y_2, \dots$  为  $\{(1, 0), (-1, 0), (0, 1), (0, -1)\}$  上的 i.i.d. 均匀分布。





刘天任

CS Theory PhD (liutianren.com)

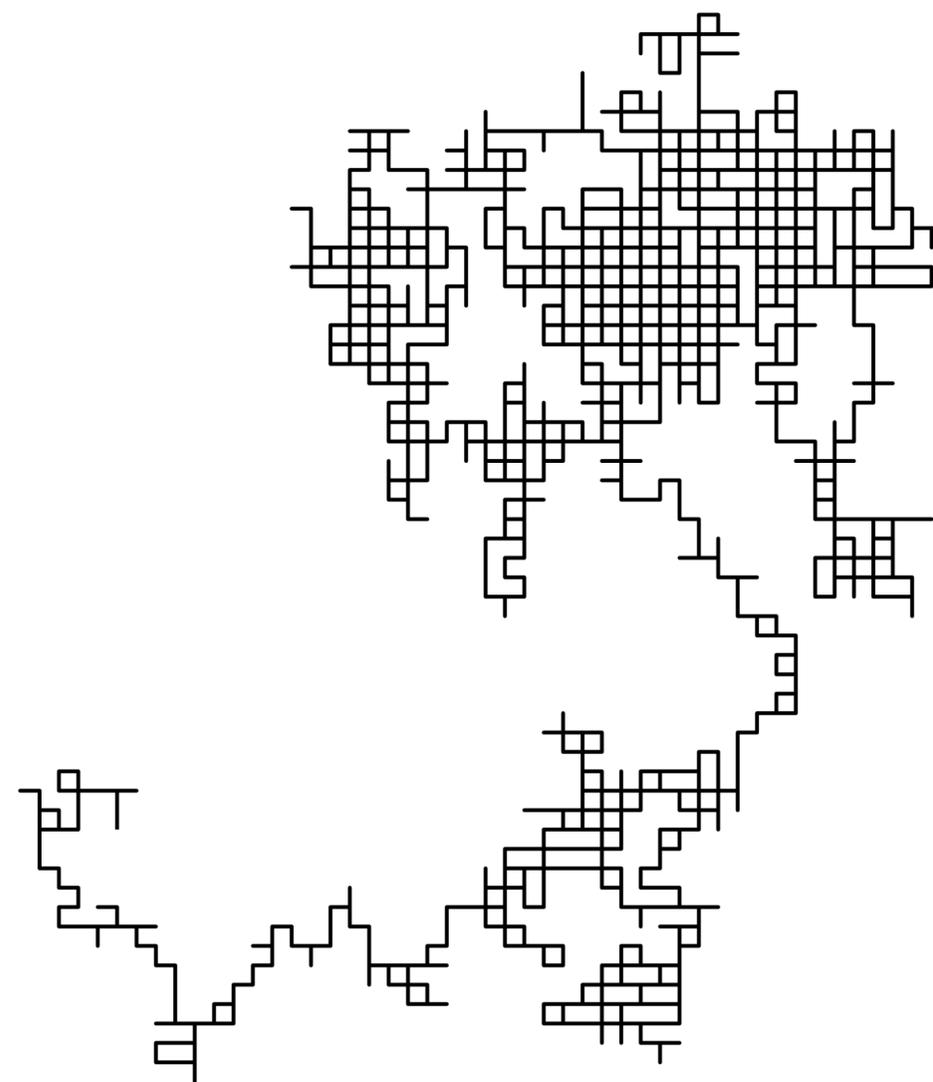
Heng Guo、Chao Xu 等 526 人赞同了该回答

假设你和女朋友都是（完美的）无头苍蝇，且商场空间有对称性的话，两个人都随机乱逛比其中一个踮着更好。

@Mather King 用平面上的布朗运动为模型。结论是双方均随机运动可使相遇时间减半。

我来补充个离散建模，也能得到同样的结果。

用二维网格上的随机漫步（random walk）来模拟随机乱逛的人。记男友在  $t$  时刻位于坐标为  $X_t$ ，女友  $t$  时刻位于坐标  $Y_t$ 。不失一般性，设  $X_0 = (0, 0)$ 。每过单位时间，两人都随机运动到旁边的一个整点。严格地说，记  $\Delta X_t := X_t - X_{t-1}$ ， $\Delta Y_t := Y_t - Y_{t-1}$ ，则  $\Delta X_1, \Delta X_2, \dots, \Delta Y_1, \Delta Y_2, \dots$  为  $\{(1, 0), (-1, 0), (0, 1), (0, -1)\}$  上的 i.i.d. 均匀分布。



随机游走：离散时间马尔可夫链





Mather King 

数学话题下的优秀回答者

2,390 人赞同了该回答

离散空间，尤其是离散时间的模型，处理这种问题特别麻烦，不如考虑一个连续模型。

商场看作 $\mathbb{R}^2$ ，两个人初始分别在 $(0,0)$ ,  $(1,0)$ . 其中一个人做标准布朗运动，或者两个人分别做独立的标准布朗运动。当两人的距离小于某个小常数，比如0.01时，视为相遇。由于二维布朗运动是邻域常返的，两人一定能相遇。

如果一个人动，另一个不动，二者的坐标差的变化就是一个标准布朗运动。如果两个人都动，坐标差的变化是两个独立标准布朗运动的和，相当于标准布朗运动的方差加倍。这等效于一个人以两倍的速度动，另一个不动，相遇用时自然减半。

简言之，一个一倍速的布朗运动，观测另一个独立的一倍速的布朗运动，看到的其实是个二倍速的布朗运动。

总之结论是随机乱逛更好，能节约一半的时间。



Mather King



数学话题下的优秀回答者

2,390 人赞同了该回答

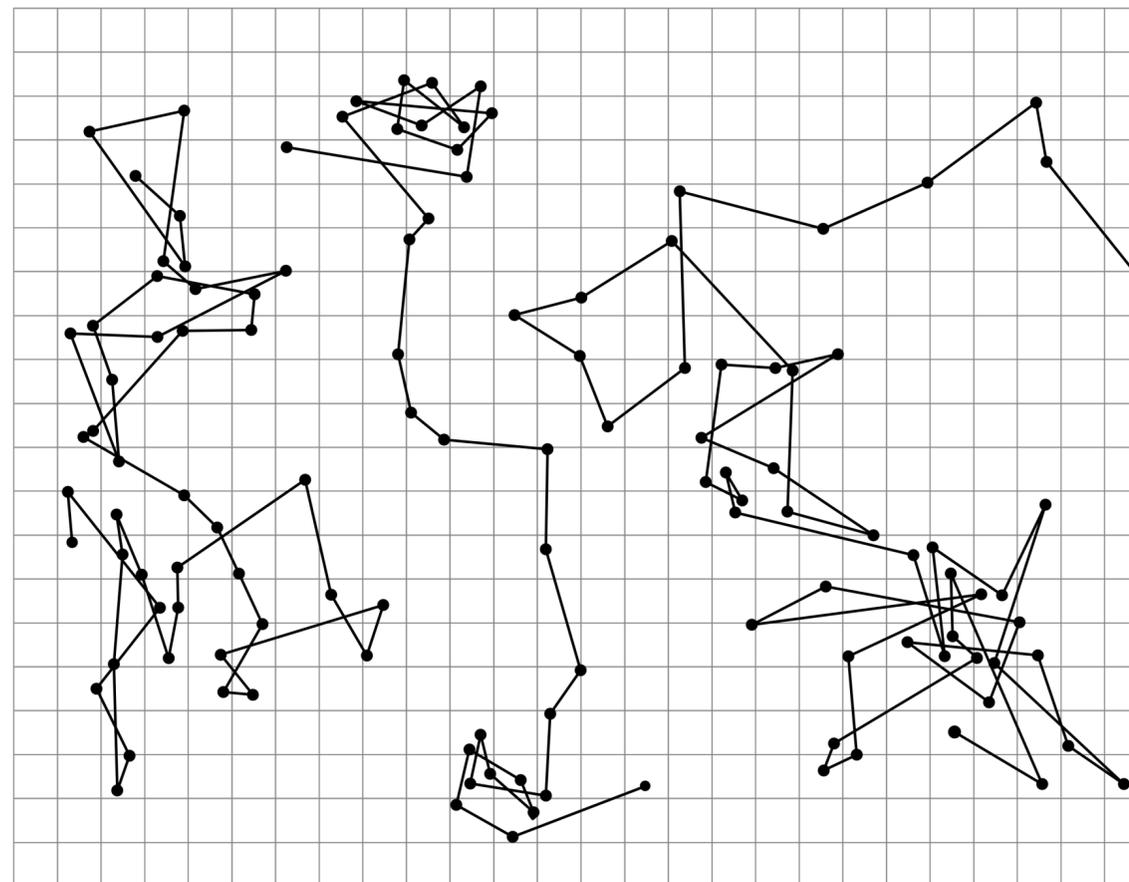
离散空间，尤其是离散时间的模型，处理这种问题特别麻烦，不如考虑一个连续模型。

商场看作 $\mathbb{R}^2$ ，两个人初始分别在 $(0,0)$ ,  $(1,0)$ . 其中一个人做标准布朗运动，或者两个人分别做独立的标准布朗运动。当两人的距离小于某个小常数，比如0.01时，视为相遇。由于二维布朗运动是邻域常返的，两人一定能相遇。

如果一个人动，另一个不动，二者的坐标差的变化就是一个标准布朗运动。如果两个人都动，坐标差的变化是两个独立标准布朗运动的和，相当于标准布朗运动的方差加倍。这等效于一个人以两倍的速度动，另一个不动，相遇用时自然减半。

简言之，一个一倍速的布朗运动，观测另一个独立的一倍速的布朗运动，看到的其实是个二倍速的布朗运动。

总之结论是随机乱逛更好，能节约一半的时间。





Mather King



数学话题下的优秀回答者

2,390 人赞同了该回答

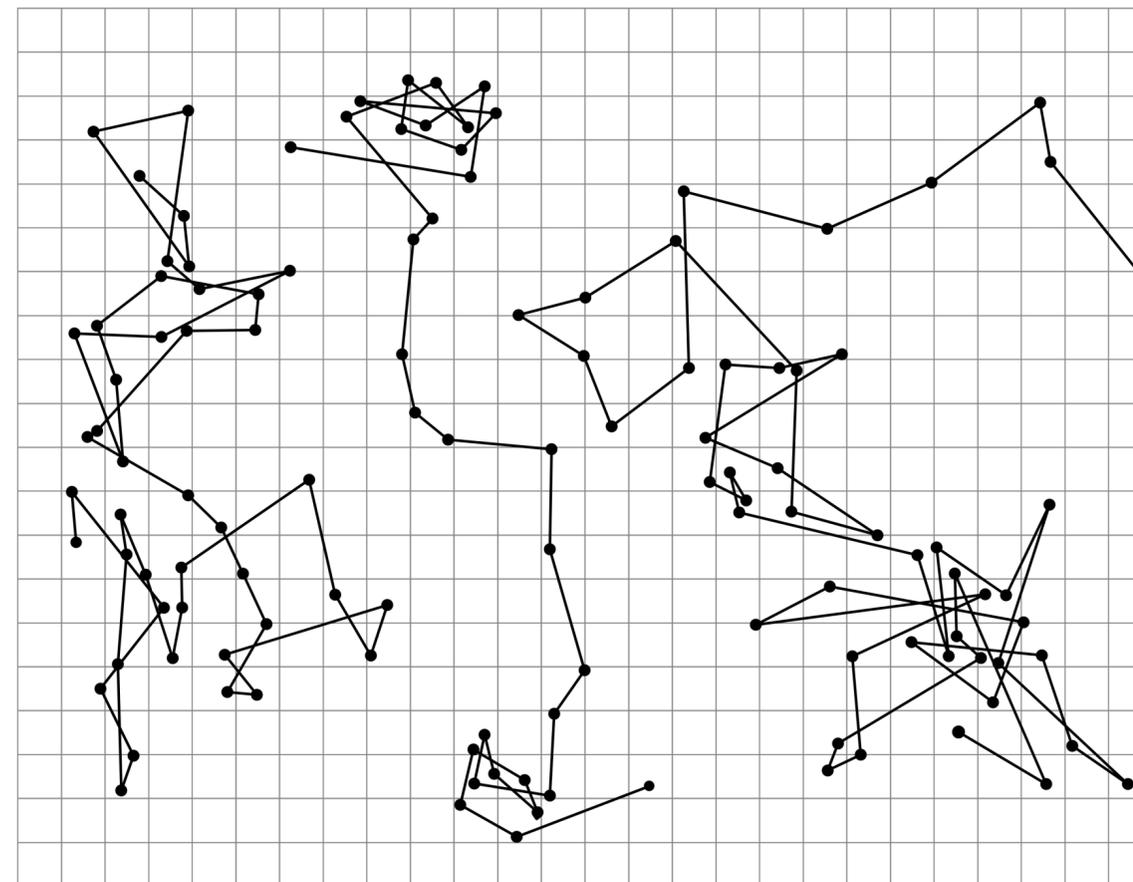
离散空间，尤其是离散时间的模型，处理这种问题特别麻烦，不如考虑一个连续模型。

商场看作 $\mathbb{R}^2$ ，两个人初始分别在 $(0,0)$ ,  $(1,0)$ . 其中一个人做标准布朗运动，或者两个人分别做独立的标准布朗运动。当两人的距离小于某个小常数，比如0.01时，视为相遇。由于二维布朗运动是邻域常返的，两人一定能相遇。

如果一个人动，另一个不动，二者的坐标差的变化就是一个标准布朗运动。如果两个人都动，坐标差的变化是两个独立标准布朗运动的和，相当于标准布朗运动的方差加倍。这等效于一个人以两倍的速度动，另一个不动，相遇用时自然减半。

简言之，一个一倍速的布朗运动，观测另一个独立的一倍速的布朗运动，看到的其实是个二倍速的布朗运动。

总之结论是随机乱逛更好，能节约一半的时间。



布朗运动 (Brownian Motion)

# 一些机器学习中的应用

# 猜天气

# 猜天气

天气矩阵  $A$ :

	晴天	阴天	下雨
晴天	$a_{11}$	$a_{12}$	$a_{13}$
阴天	$a_{21}$	$a_{22}$	$a_{23}$
下雨	$a_{31}$	$a_{32}$	$a_{33}$

# 猜天气

天气矩阵  $A$ :

	晴天	阴天	下雨
晴天	$a_{11}$	$a_{12}$	$a_{13}$
阴天	$a_{21}$	$a_{22}$	$a_{23}$
下雨	$a_{31}$	$a_{32}$	$a_{33}$

行为矩阵  $B$ :

	宅	出门
晴天	$b_{11}$	$b_{12}$
阴天	$b_{21}$	$b_{22}$
下雨	$b_{31}$	$b_{32}$

# 猜天气

天气矩阵  $A$ :

	晴天	阴天	下雨
晴天	$a_{11}$	$a_{12}$	$a_{13}$
阴天	$a_{21}$	$a_{22}$	$a_{23}$
下雨	$a_{31}$	$a_{32}$	$a_{33}$

行为矩阵  $B$ :

	宅	出门
晴天	$b_{11}$	$b_{12}$
阴天	$b_{21}$	$b_{22}$
下雨	$b_{31}$	$b_{32}$

观测行为:

宅出出宅出出宅出宅宅出...

# 猜天气

天气矩阵  $A$ :

	晴天	阴天	下雨
晴天	$a_{11}$	$a_{12}$	$a_{13}$
阴天	$a_{21}$	$a_{22}$	$a_{23}$
下雨	$a_{31}$	$a_{32}$	$a_{33}$

观测行为:

宅出出宅出出宅出宅出...

行为矩阵  $B$ :

	宅	出门
晴天	$b_{11}$	$b_{12}$
阴天	$b_{21}$	$b_{22}$
下雨	$b_{31}$	$b_{32}$

猜测  $A$ 、 $B$  与第一天的天气分布

# 猜天气

天气矩阵  $A$ :

	晴天	阴天	下雨
晴天	$a_{11}$	$a_{12}$	$a_{13}$
阴天	$a_{21}$	$a_{22}$	$a_{23}$
下雨	$a_{31}$	$a_{32}$	$a_{33}$

观测行为:

宅出出宅出出宅出宅宅出...

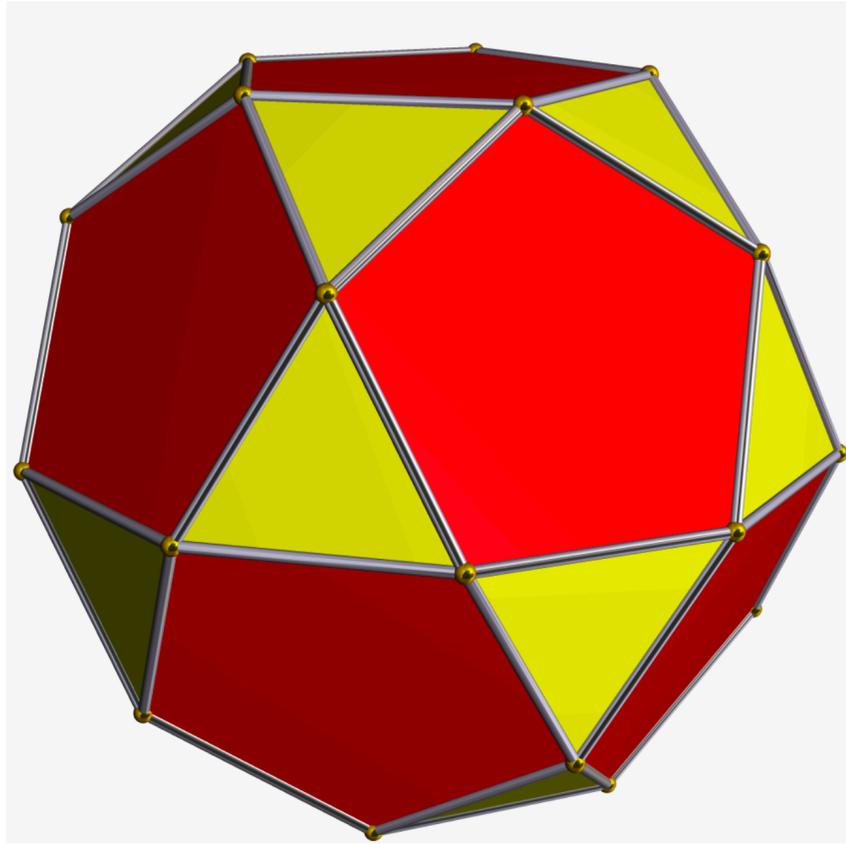
隐马尔可夫模型 (Hidden Markov Model)

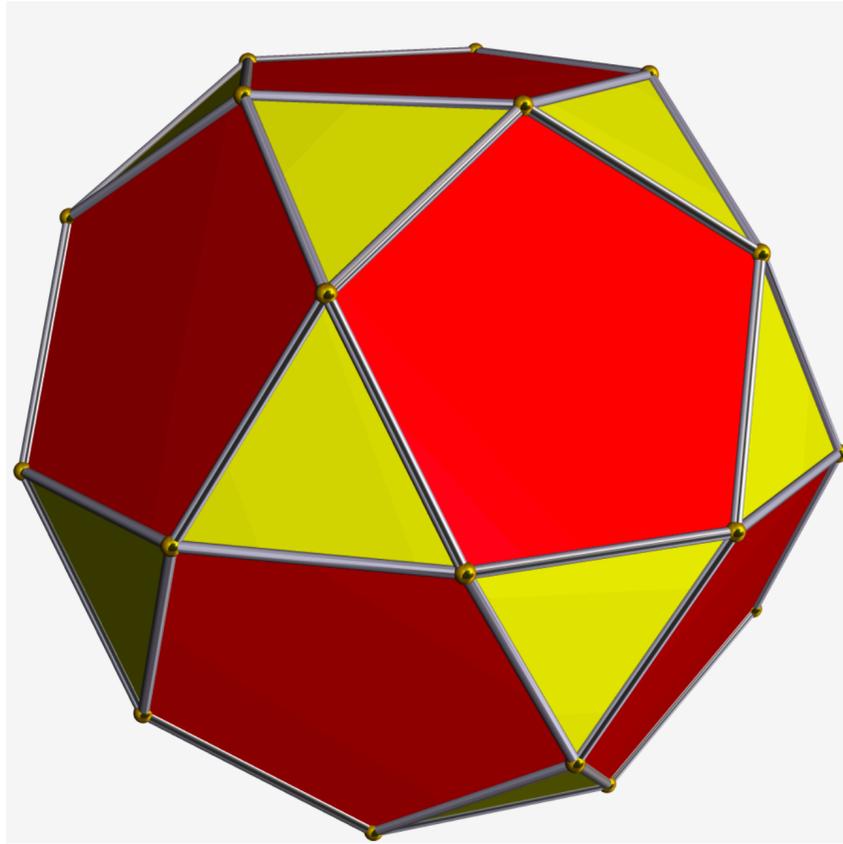
行为矩阵  $B$ :

	宅	出门
晴天	$b_{11}$	$b_{12}$
阴天	$b_{21}$	$b_{22}$
下雨	$b_{31}$	$b_{32}$

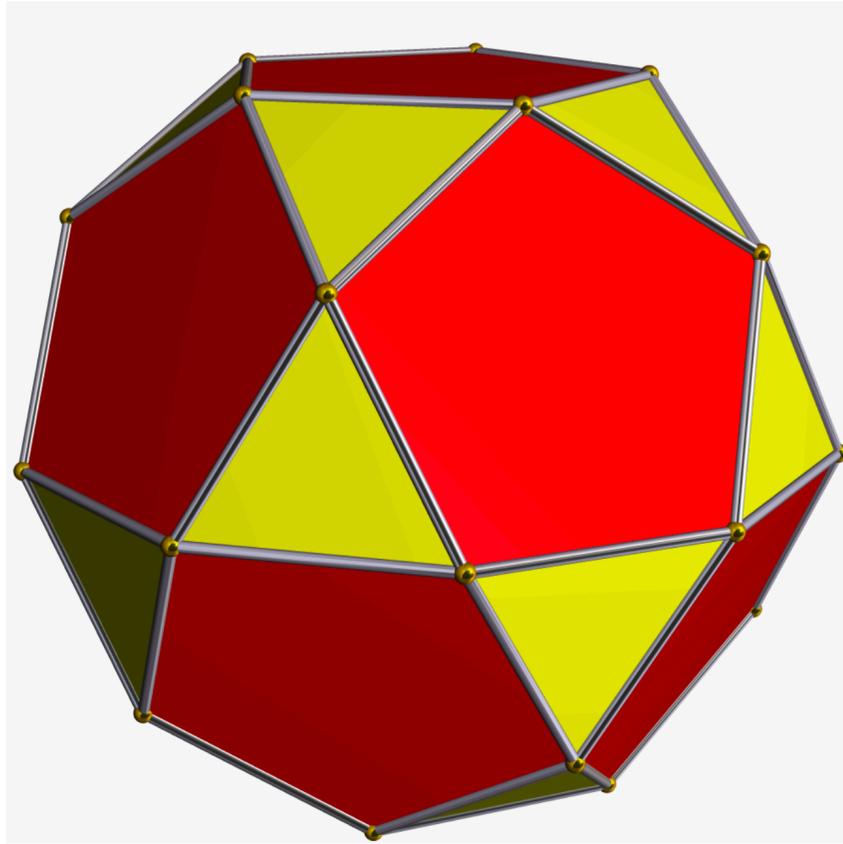
猜测  $A$ 、 $B$  与第一天的天气分布





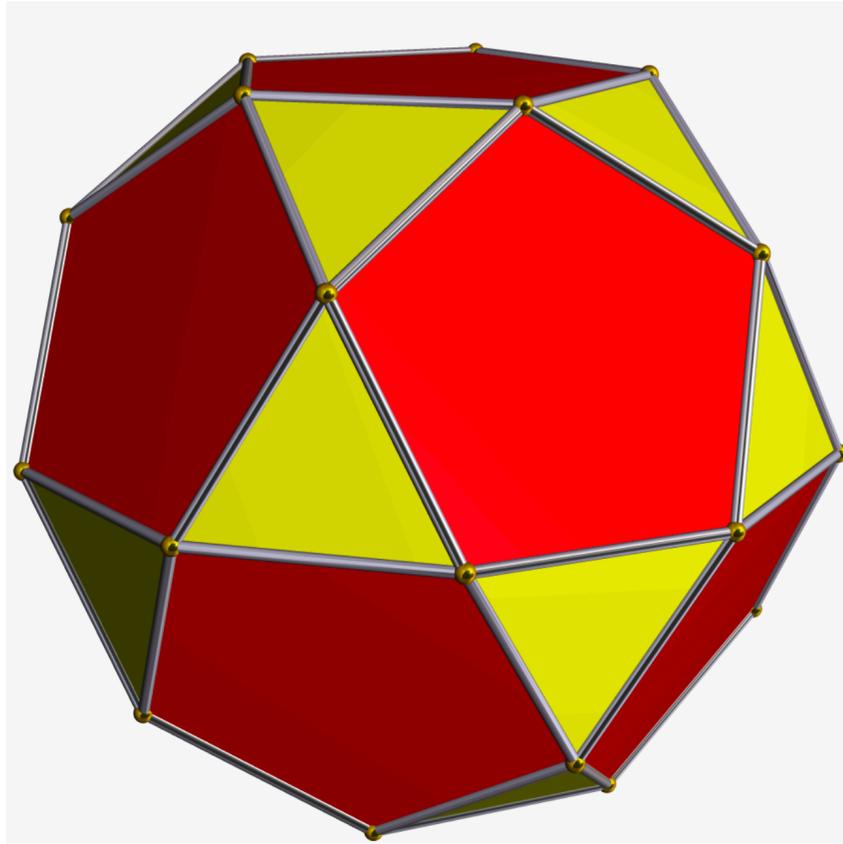


给定一个凸集  $P$ ，和一个  
membership oracle



给定一个凸集  $P$ ，和一个 membership oracle

$$\forall x \in \mathbb{R}^n, \quad f(x) = \begin{cases} 1, & \text{if } x \in P \\ 0, & \text{if } x \notin P \end{cases}$$



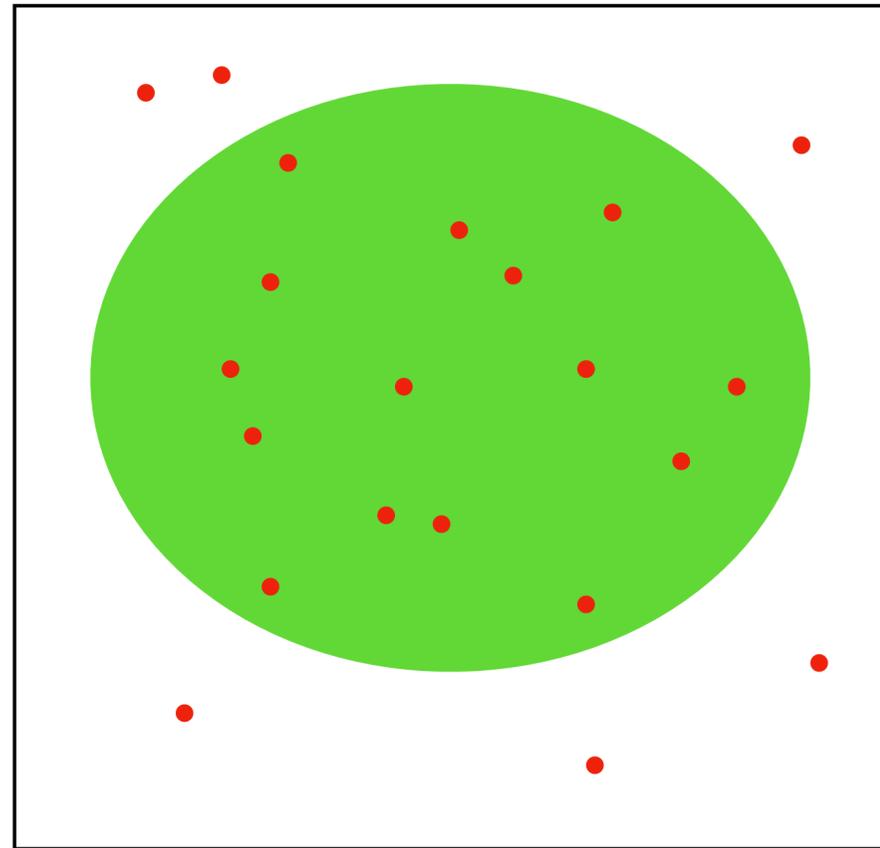
给定一个凸集  $P$ ，和一个  
membership oracle

$$\forall x \in \mathbb{R}^n, \quad f(x) = \begin{cases} 1, & \text{if } x \in P \\ 0, & \text{if } x \notin P \end{cases}$$

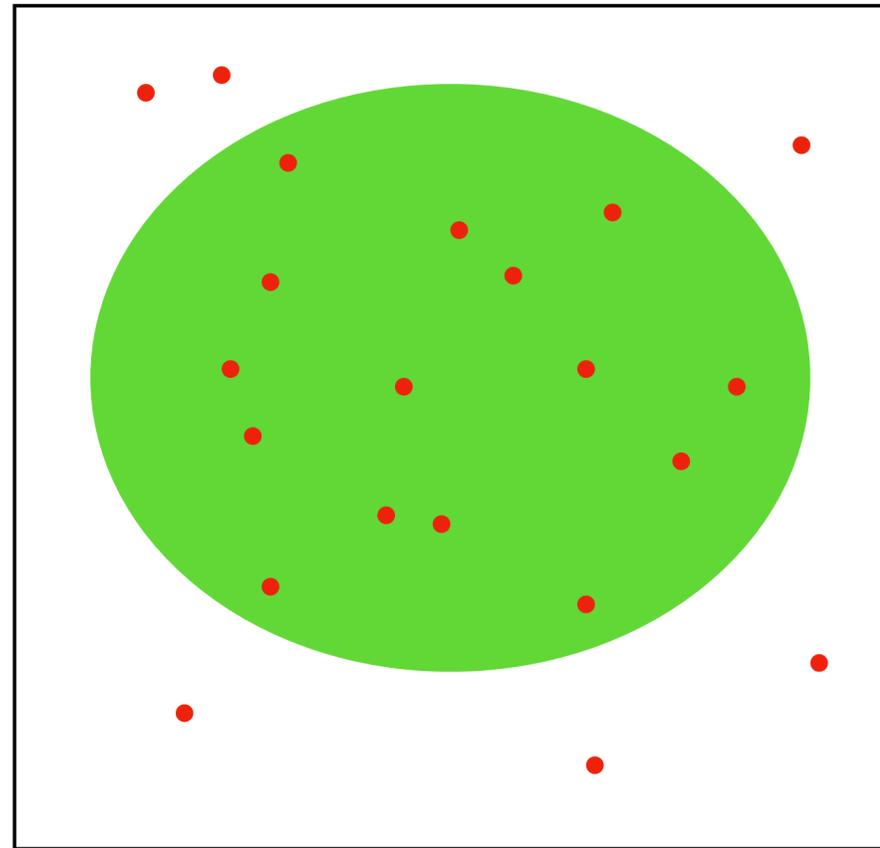
如何计算  $P$  的体积?

# 方法一：蒙特卡洛法

# 方法一：蒙特卡洛法

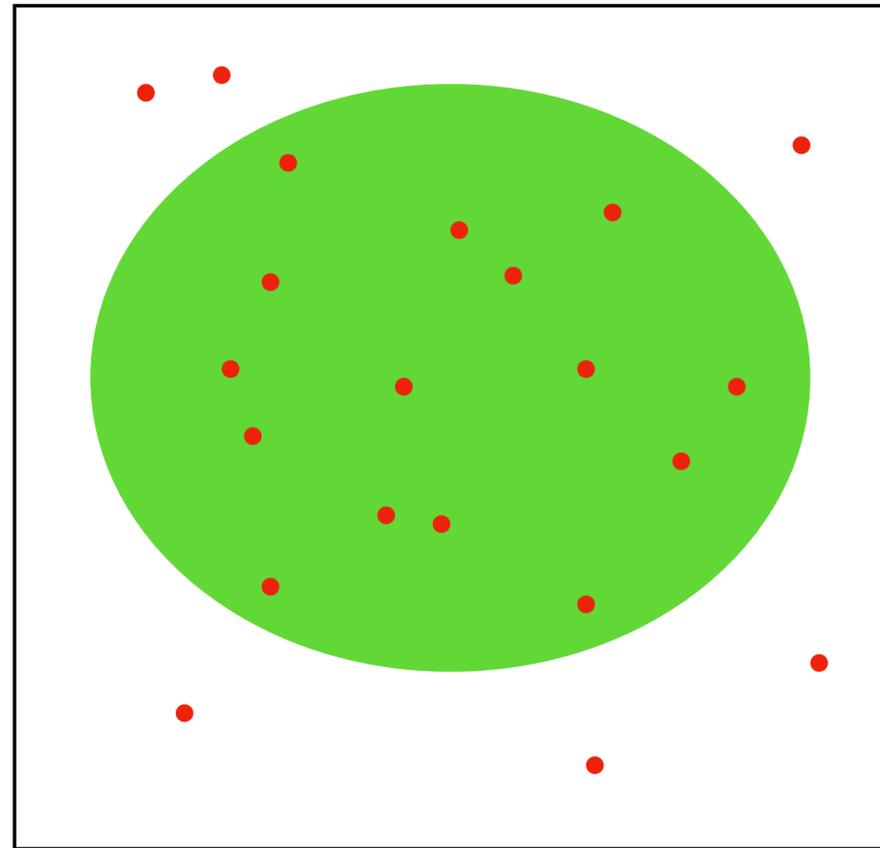


# 方法一：蒙特卡洛法



高维诅咒 (curse of dimensionality) :

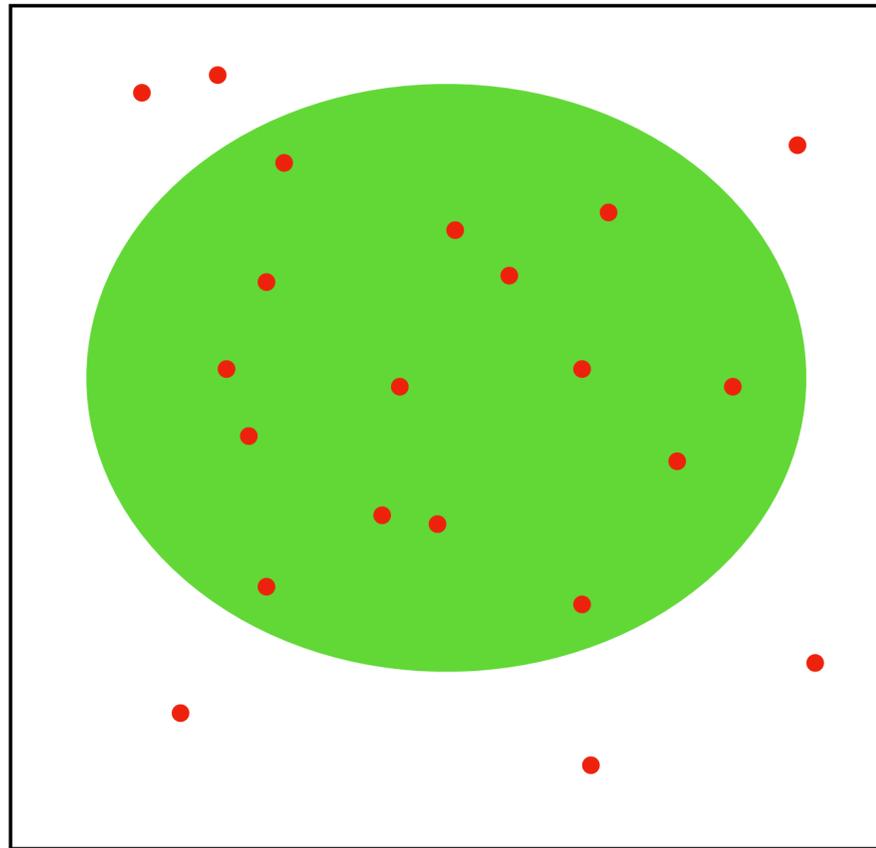
# 方法一：蒙特卡洛法



高维诅咒 (curse of dimensionality) :

$$n \text{ 维单位球体积} = O(R^n)$$

# 方法一：蒙特卡洛法



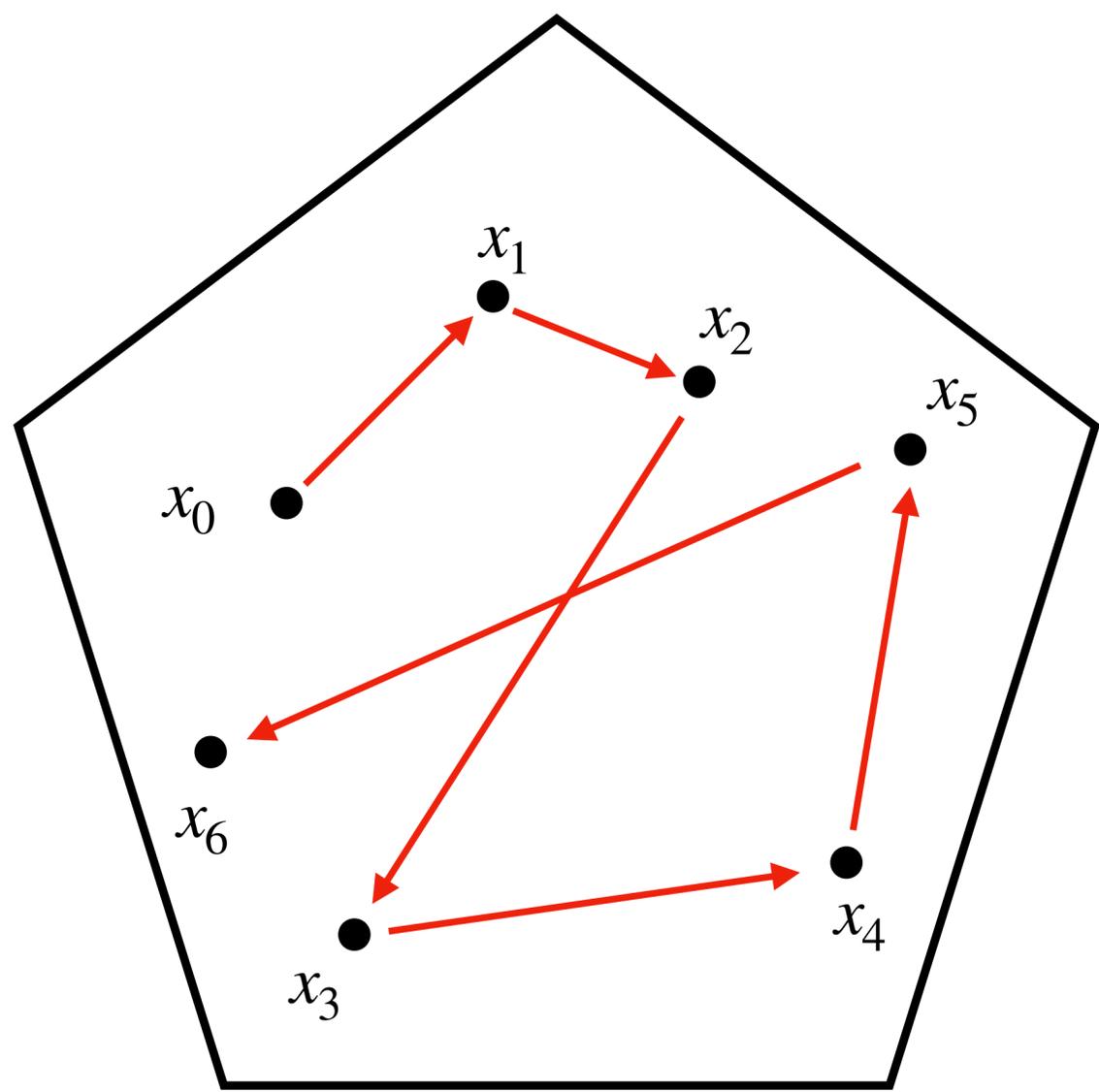
高维诅咒 (curse of dimensionality) :

$n$  维单位球体积  $= O(R^n)$

需要至少  $2^{\Omega(n)}$  个样本

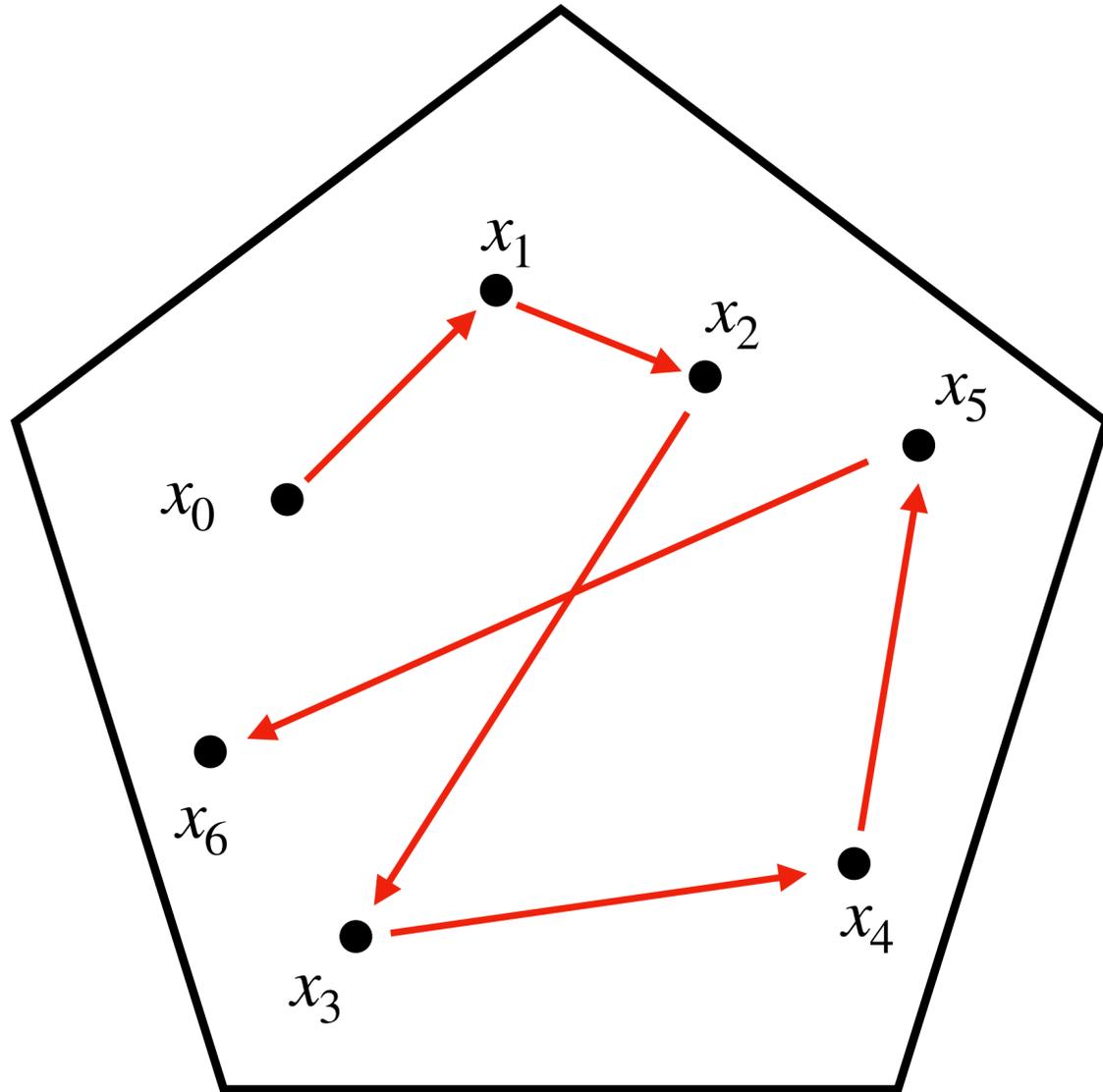
# 方法二：马尔可夫链蒙特卡洛法

# 方法二：马尔可夫链蒙特卡洛法

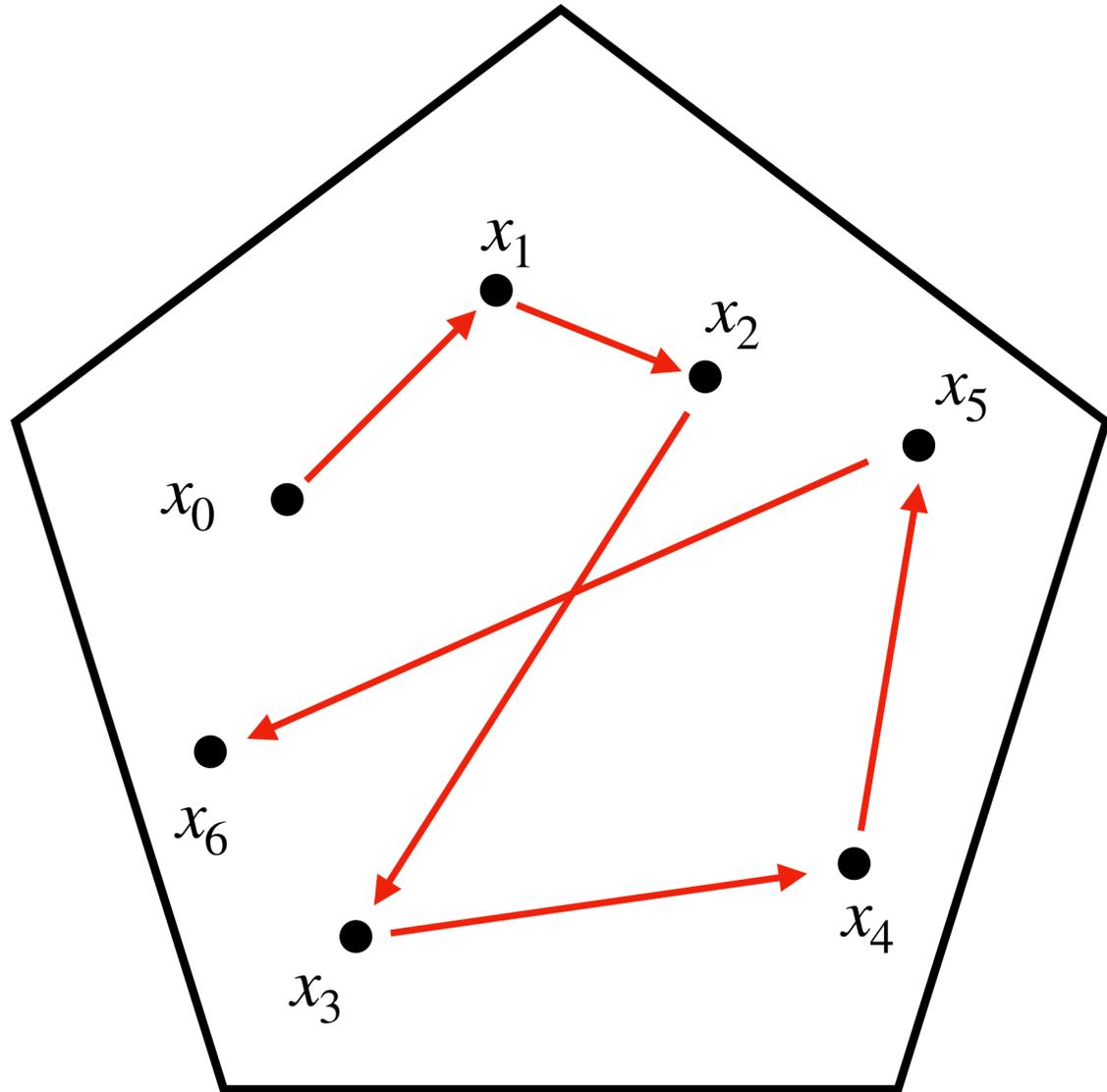


# 方法二：马尔可夫链蒙特卡洛法

Ball Walk



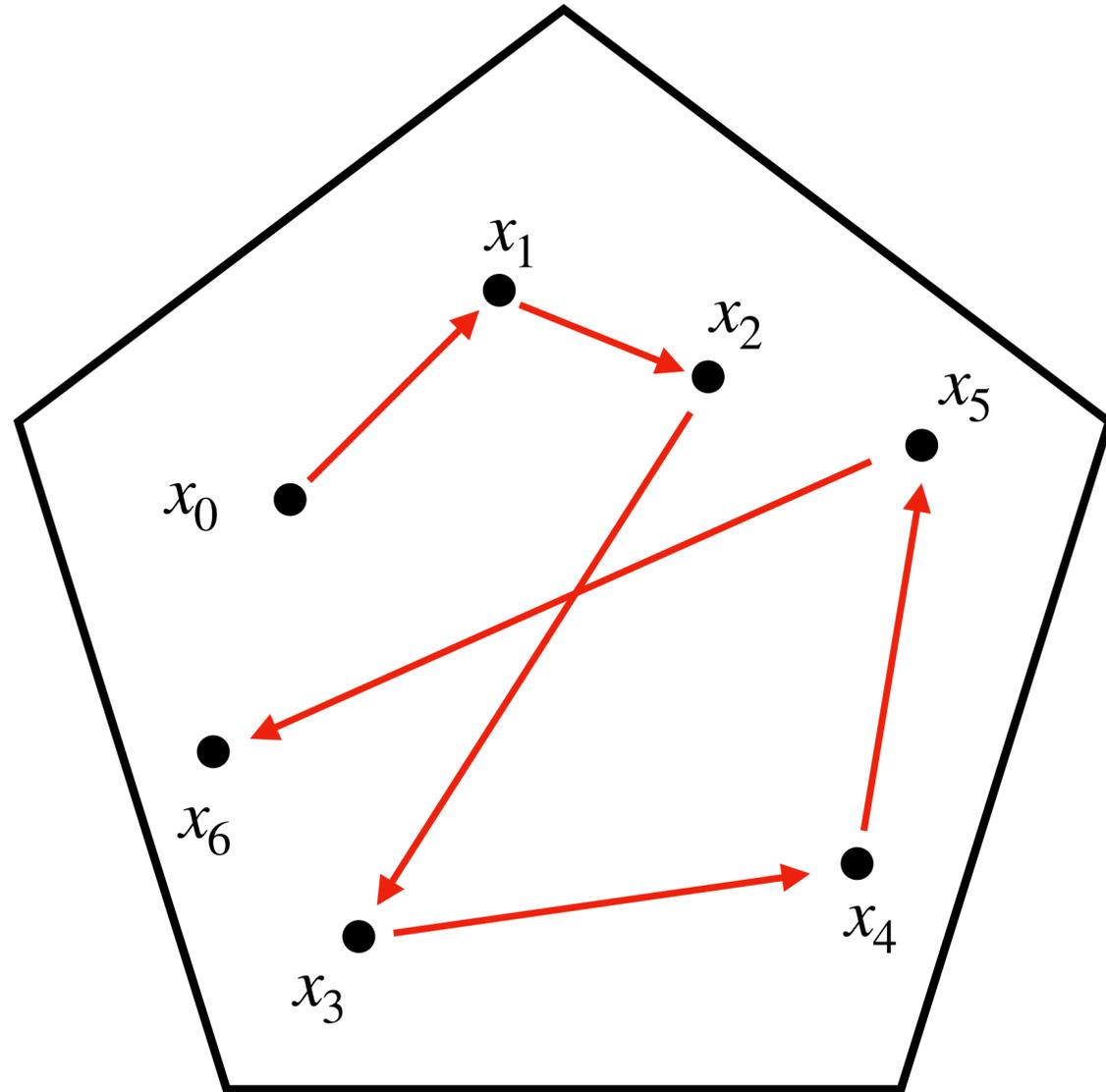
# 方法二：马尔可夫链蒙特卡洛法



Ball Walk

$$x_{t+1} = x_t + R_t, \quad R_t \in_R \text{Ball}(0, r)$$

# 方法二：马尔可夫链蒙特卡洛法

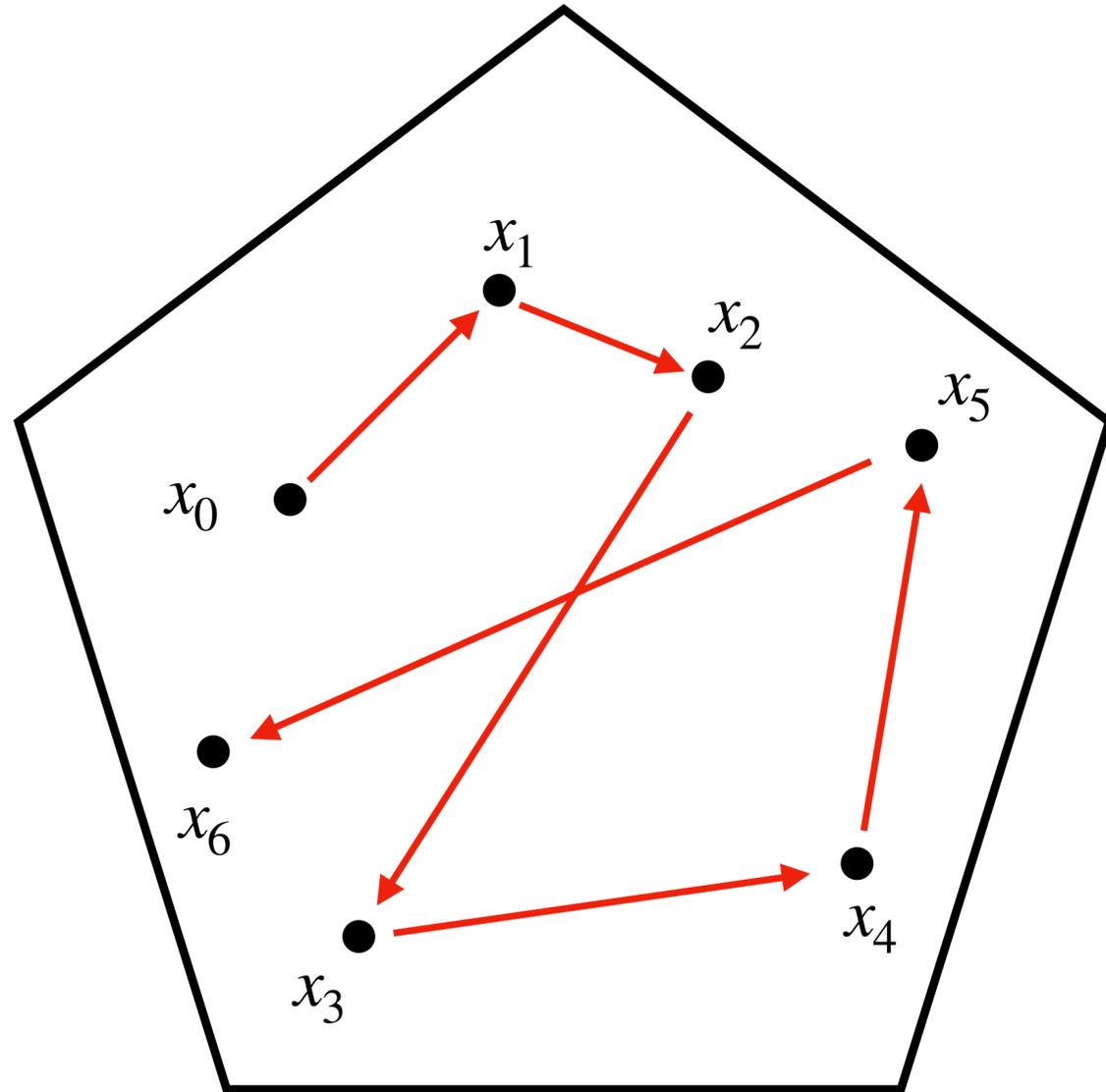


Ball Walk

$$x_{t+1} = x_t + R_t, \quad R_t \in_R \text{Ball}(0, r)$$

多项式步收敛！

# 方法二：马尔可夫链蒙特卡洛法



Ball Walk

$$x_{t+1} = x_t + R_t, \quad R_t \in_R \text{Ball}(0, r)$$

多项式步收敛！

马尔可夫过程 (Markov Process)

# AI2613是

- 马尔可夫链（离散/连续）
- 泊松过程
- 马尔可夫过程
- 布朗运动
- ...

传统随机过程

+

AI/大数据/机器学习中的算法应用