

第3章 随机游动

设 G 是一个有限图. 考虑在 G 的顶点上做如下类型的游动. 从顶点 u (顶点 u 可以根据某一概率分布随机选取或预先指定) 出发, 在与 u 相关联的所有的边中, 均匀地随机选取一条 (即如果有 k 条边与 u 相连, 则每条边被选取的概率都是 $1/k$). 移动到被选边的另一顶点 v 上, 然后从 v 继续之前的操作. 熟悉概率论的读者会发现这种随机游动是有限状态 Markov 链的一个特例. 关于这种游动有很多有趣的问题可以问; 最基本的一个问题就是计算在给定的 ℓ 步后到达一个给定顶点的概率.

假设顶点 u 的度 (degree) 为 d_u , 即有 d_u 条边与 u 关联 (u 的每个自环仅计算一次). 记 $M = M(G)$ 为行与列都按 G 的顶点集 $\{v_1, \dots, v_p\}$ 依次标记的矩阵, 其 (u, v) -元为

$$M_{uv} = \frac{\mu_{uv}}{d_u}, \quad (3.1)$$

其中 μ_{uv} 为 u 与 v 之间边的数目 (对简单图这个数为 0 或 1). 因而 M_{uv} 正好是从 u 出发, 下一步到达 v 的概率. 称 M 为与 G 相关联的概率矩阵 (probability matrix). 经过简单的概率论证 (等价于定理 1.1) 可以证明, 如果 ℓ 是一个正整数, 则 $(M^\ell)_{uv}$ 等于从 u 出发经过 ℓ 步到达 v 的概率. 现在假设起点不是指定的, 而是给定和为 1 的概率值 ρ_u , 并以概率 ρ_u 从顶点 u 出发. 令 P 为行向量 $P = [\rho_{v_1}, \dots, \rho_{v_p}]$. 同样, 通过初等的论证可以证明, 如果 $PM^\ell = [\sigma_{v_1}, \dots, \sigma_{v_p}]$, 那么 σ_v 是 ℓ 步后到达 v 的概率 (有给定的起始分布). 如第 1 章中的推导, 若得到了 M 的特征值和特征向量, 那么我们就可以计算至关重要的概率 $(M^\ell)_{uv}$ 和 σ_u .

既然矩阵 M 和邻接矩阵 A 不一样, 那么以上这些和邻接矩阵又有什么关系呢? 答案是在一个重要的情况下, M 正好是 A 的数量积. 如果每一个 $d_u = d$, 则称图 G 是度为 d 的正则图 (regular graph), 即每个顶点都与 d 条边关联. 在这种情形下很容易看到 $M(G) = \frac{1}{d}A(G)$. 因此 $M(G)$ 和 $A(G)$ 的特征向量 E_u 相同, 特征值为 $\lambda_u(M) = \frac{1}{d}\lambda_u(A)$. 这样在正则图上的随机游动就与图的邻接矩阵紧密相关.

3.1 例 考虑在 n 维立方体 C_n 上的一个随机游动, 起点在“原点”(向量 $(0, \dots, 0)$). 问经过 ℓ 步后再次回到原点的概率 p_ℓ 是多少? 在用任何公式之前, 注意到在偶数 (奇数) 步之后, 游动肯定到达了有偶数 (奇数) 个 1 的顶点上. 因此当 ℓ 是奇数时 $p_\ell = 0$. 现在注意到

C_n 是度为 n 的正则图. 因此由 (2.6) 有

$$\lambda_u(\mathbf{M}(C_n)) = \frac{1}{n}(n - 2\omega(u)).$$

再根据 (2.10) 我们就能得到

$$p_\ell = \frac{1}{2^n n^\ell} \sum_{i=0}^n \binom{n}{i} (n - 2i)^\ell.$$

注意到在上式中, 当 ℓ 是奇数时 p_ℓ 确实化简为 0.

值得注意的是尽管概率矩阵 \mathbf{M} 不一定是对称矩阵, 但它仅有实特征值.

3.2 定理 设 G 是一个有限图. 那么概率矩阵 $\mathbf{M} = \mathbf{M}(G)$ 可对角化且仅有实特征值.

证明 可以假设 G 是连通的且至少有两个顶点, 这样对 G 中的每一个顶点 v 都有 $d_v > 0$.

令 \mathbf{D} 为行与列都按 G 的顶点标记的对角矩阵, 且 $D_{vv} = \sqrt{d_v}$. 那么

$$\begin{aligned} (\mathbf{DMD}^{-1})_{uv} &= \sqrt{d_u} \cdot \frac{\mu_{uv}}{d_u} \cdot \frac{1}{\sqrt{d_v}} \\ &= \frac{\mu_{uv}}{\sqrt{d_u d_v}}. \end{aligned}$$

因此 \mathbf{DMD}^{-1} 是对称矩阵, 故仅有实特征值. 而如果 \mathbf{B} 和 \mathbf{C} 都是任意的 $p \times p$ 矩阵且 \mathbf{C} 是可逆的, 则 \mathbf{B} 和 \mathbf{CBC}^{-1} 有相同的特征多项式, 因此有相同的特征值. 这样 \mathbf{M} 的特征值都是实数. 此外, \mathbf{B} 可对角化当且仅当 \mathbf{CBC}^{-1} 可对角化. (事实上, \mathbf{B} 和 \mathbf{CBC}^{-1} 有相同的 Jordan 标准型.) 因为对称矩阵可对角化, 所以 \mathbf{M} 可对角化. \square

我们给出一个更进一步的例子说明线性代数和图上的随机游动之间的联系. 设 u 和 v 是连通图 G 的顶点. 定义访问时间 (access time) 或者到达时间 (hitting time) $H(u, v)$ 为在如上定义的随机游动中, 自 u 出发首次到达 v 所走步数的期望值. 因此如果走 n 步后首次到达 v 的概率为 p_n , 则由期望的定义可得

$$H(u, v) = \sum_{n \geq 1} np_n. \quad (3.2)$$

可以想象这个和式可能趋于无穷, 然而下文中我们将会看到这种情形不会出现. 注意到 $H(v, v) = 0$.

举个例子, 假设 G 有三个顶点 u, v, w , 在 u 和 w 之间有一条边, 在 w 与 v 之间有另一条边. 我们可以如下这样计算 $H(u, v)$. 第一步将到达 w . 然后以 $\frac{1}{2}$ 的概率到达 v 和以 $\frac{1}{2}$ 的概率回到 u . 因此 [为什么]

$$H(u, v) = \frac{1}{2} \cdot 2 + \frac{1}{2}(2 + H(u, v)). \quad (3.3)$$

解这个线性方程可得 $H(u, v) = 4$.

我们希望给出一个用线性代数计算到达时间 $H(u, v)$ 的公式. 证明中需要非负矩阵特征值和特征向量的一些基本结果, 我们将陈述和解释这些结果, 但不提供证明. 一个 $r \times r$ 实矩阵 B , 当其所有元都非负时, 则称它为非负的 (nonnegative). 当 B 不是 1×1 矩阵 [0] 且不存在一个置换矩阵 P (每行每列都恰有一个 1, 且其他元都为 0 的矩阵) 使得

$$PBP^{-1} = \begin{bmatrix} C & D \\ \mathbf{0} & E \end{bmatrix},$$

其中 C 和 E 是阶数大于 0 的方阵, 则称 B 为不可约的 (irreducible). 例如, 图 G 的邻接矩阵 A 和概率矩阵 M 是不可约的当且仅当 G 是连通的且不是一个孤立顶点 (isolated vertex) (也就是说, 顶点 v 没有关联任何边, 甚至没有 v 到它本身的自环). 我们现在给出未加证明的某一版本的 Perron-Frobenius 定理. Perron-Frobenius 定理中还有些其他部分但这里不需要就省略了.

3.3 定理 设 B 是非负不可约的方阵. 若 ρ 是 B 所有特征值的绝对值的最大值, 则 $\rho > 0$, 并且存在一个特征值等于 ρ . 此外, 存在对应于 ρ 的所有元均为正数的特征向量 (该特征向量在相差一个正实数倍的意义下是唯一的).

现在令 M 为等式 (3.1) 定义的概率矩阵. 用 $M[v]$ 表示 M 中删除以 v 标记的那一行和那一列后余下的矩阵. 因此若 G 有 p 个顶点, 则 $M[v]$ 是一个 $(p-1) \times (p-1)$ 的矩阵. 令 $T[v]$ 是 $p-1$ 维列向量, 它的行按顶点 $w \neq v$ 标记, 且 $T[v]_w = \mu_{wv}/d_w$. 记 I_{p-1} 为 $p-1$ 阶单位矩阵.

3.4 定理 矩阵 $I_{p-1} - M[v]$ 是可逆的, 且

$$H(u, v) = ((I_{p-1} - M[v])^{-2}T[v])_u, \quad (3.4)$$

等式右端为列向量 $(I_{p-1} - M[v])^{-2}T[v]$ 中与顶点 u 对应的元.

证明 我们首先给出一个“形式上”的论证然后再验证其正确性. 从 u 出发, 从未经过 v , n 步后到达某一顶点 w 的概率为 $(M[v]^n)_{uw}$ [为什么]. 一旦到达了 w 下一步到 v 的概率是 μ_{wv}/d_w . 因此由期望的定义可得

$$H(u, v) = \sum_{w \neq v} \sum_{n \geq 0} (n+1) \frac{\mu_{wv}}{d_w} (M[v]^n)_{uw}. \quad (3.5)$$

我们断言: 若 x 是满足 $|x| < 1$ 的复数, 则

$$\sum_{n \geq 0} (n+1)x^n = (1-x)^{-2}. \quad (3.6)$$

上式是微积分中一个简单的练习. 例如, 我们可以计算乘积 $(1-x)^2 \sum_{n \geq 0} (n+1)x^n$ 中 x^n 的系数. 也可以对下面这个熟悉的等式进行求导

$$\sum_{n \geq 0} x^n = \frac{1}{1-x}. \quad (3.7)$$

另一种证明方法是应用二项式定理对指数为 -2 的情形将 $(1-x)^{-2}$ 展开. $|x| < 1$ 时收敛的结论可由例如等式 (3.7) 成立的条件得到.

我们“盲目地”将等式 (3.6) 应用到等式 (3.5) 中, 可得

$$\begin{aligned} H(u, v) &= \sum_{w \neq v} ((I_{p-1} - M[v])^{-2})_{uw} \frac{\mu_{wv}}{d_w} \\ &= ((I_{p-1} - M[v])^{-2} T[v])_u, \end{aligned} \quad (3.8)$$

得证.

剩下的就是验证等式 (3.8) 中的推导. 对任意的实 (或复) $r \times r$ 矩阵 B , 可以逐项定义 $\sum_{n \geq 0} (n+1)B^n$, 也就是说, 若对 B 和 C 的所有行列标号 i 和 j 都有

$$\sum_{n \geq 0} (n+1)(B^n)_{ij} = C_{ij},$$

则记 $\sum_{n \geq 0} (n+1)B^n = C$.

通过对 m 的归纳容易证明如下等式

$$(I_r - B)^2 (I_r + 2B + 3B^2 + \cdots + mB^{m-1}) = I_r - (m+1)B^m + mB^{m+1}. \quad (3.9)$$

假设 B 是可对角化的且 B 所有的特征值 $\lambda_1, \dots, \lambda_r$ 都满足 $|\lambda_j| < 1$. 注意到等式 (1.1) 的证明可以扩展到任意可对角化的矩阵. (矩阵 U 没必要是正交的, 但这与当前的证明无关.) 因此

$$(B^n)_{ij} = c_1 \lambda_1^n + \cdots + c_r \lambda_r^n,$$

其中 c_1, \dots, c_r 都是 (与 n 无关的) 复数. 可以看到当 $m \rightarrow \infty$ 时等式 (3.9) 右端的极限为 I_r . 这样 [为什么] $\sum_{n \geq 0} (n+1)B^n$ 收敛于 $(I_r - B)^{-2}$.

注 上面的讨论证明了 $I_r - B$ 的确是可逆的. 这也可以很快地从 B 的所有特征值的绝对值都小于 1 的假设中得到, 因为特别地, $\lambda = 1$ 不是特征值.

通过上面的讨论可知, 剩下需要证明 $M[v]$ 是可对角化的, 且它的所有特征值的绝对值都小于 1. $M[v]$ 的对角化与定理 3.2 中证明 M 的对角化完全一样. (因此我们同样可得 $M[v]$ 只有实特征值, 然而这里不需要这一结论.) 剩下证明 $M[v]$ 的特征值 $\theta_1, \dots, \theta_{p-1}$ 满

足 $|\theta_j| < 1$. 我们想对矩阵 $M[v]$ 应用定理 3.3, 但是这个矩阵有可能不是不可约的, 因为图 $G-v$ (定义为从 G 中删除顶点 v 及它的所有关联边) 未必连通, 也可能恰是一个孤立顶点. 若 $G-v$ 的连通分支为 H_1, \dots, H_m , 则可以重排 $G-v$ 顶点的顺序使得 $M[v]$ 有如下的块状结构

$$M[v] = \begin{bmatrix} N_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & N_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & N_m \end{bmatrix},$$

其中每一个 N_i 都是不可约的或者是 1×1 的矩阵 $[0]$ (对应的 H_i 是一个孤立顶点). $M[v]$ 的特征值就是 N_i 的特征值.

我们需要证明 N_i 的每个特征值的绝对值都小于 1. 若 $N_i = [0]$ 则它的特征值只有 0, 因此我们可以假设 H_i 不是一个孤立顶点. 假设 H_i 有 k 个顶点, 则 N_i 是 $k \times k$ 矩阵. 令 ρ_i 为 N_i 的最大实特征值, 那么由定理 3.3 可知 N_i 的所有特征值 λ 都满足 $|\lambda| \leq \rho_i$. 注意 N_i 与 N_i^t 有相同的特征值. 设列向量 $U = [u_1, \dots, u_k]^t$ 为 N_i^t 关于 ρ_i 的特征向量, 且它的所有元都为正数 (由定理 3.3 可知它的存在性). 用 V 表示长为 k 的全 1 行向量. 考虑矩阵的乘积 VN_i^tU . 一方面我们有

$$VN_i^tU = V(\rho_i U) = \rho_i(u_1 + \cdots + u_k). \quad (3.10)$$

另一方面, 若 σ_j 表示 N_i^t 的第 j 列各元的和, 则

$$VN_i^tU = [\sigma_1, \dots, \sigma_k]U = \sigma_1 u_1 + \cdots + \sigma_k u_k. \quad (3.11)$$

现在每一个 σ_j 都满足 $0 \leq \sigma_j \leq 1$, 且至少有一个 σ_h 满足 $\sigma_h < 1$ [为什么]. 因为每一个 $u_j > 0$, 从等式 (3.11) 可得 $VN_i^tU < u_1 + \cdots + u_k$. 再与等式 (3.10) 比较可得 $\rho_i < 1$.

因为 $M[v]$ 的特征值就是 N_i 的特征值, 这样 $M[v]$ 的所有特征值 θ 都满足 $|\theta| < 1$. 定理 3.4 证毕. \square

3.5 例 设 G 为图 3.1 所示的图并令 $v = v_4$. 那么

$$M = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 \end{bmatrix},$$

$$I_3 - M[v] = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & 0 \\ -\frac{1}{4} & 1 & -\frac{1}{4} \\ 0 & -\frac{1}{2} & 1 \end{bmatrix},$$

$$(I_3 - M[v])^{-2} = \begin{bmatrix} \frac{55}{16} & \frac{13}{6} & \frac{17}{24} \\ \frac{13}{8} & \frac{7}{3} & \frac{11}{12} \\ \frac{17}{16} & \frac{11}{6} & \frac{13}{8} \end{bmatrix},$$

$$(I_3 - M[v])^{-2} \begin{bmatrix} \frac{1}{3} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{31}{12} \\ \frac{13}{6} \\ \frac{25}{12} \end{bmatrix}.$$

因此 $H(v_1, v) = 31/12$, $H(v_2, v) = 13/6$, 和 $H(v_3, v) = 25/12$.

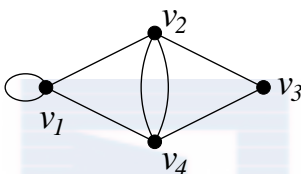


图 3.1 例 3.5 的图 G

注 用于证明当 B 的所有特征值的绝对值小于 1 时, $\sum_{n \geq 0} (n+1)B^n$ 收敛的方法可以推广, 通过进一步努力 (主要关于非对角化的情形) 可以证明如下结论. 令 $F(x) = \sum_{n \geq 0} a_n x^n$ 为一个幂级数, 其系数 a_n 为复数. 设 $\alpha > 0$ 为使得 $|x| < \alpha$ 时 $F(x)$ 收敛的数. 用 B 表示所有特征值 λ 都满足 $|\lambda| < \alpha$ 的复方阵. 那么矩阵幂级数 $\sum_{n \geq 0} a_n B^n$ 具有如前文描述的逐项收敛的性质.

第 3 章 注记

图上的随机游动是一个广泛的课题, 这里我们仅仅只接触到了一点皮毛. 有两个比我们当前考虑的问题深刻得多的典型问题: 一个随机游动以多快的速度趋于练习 3.1 中所定义的平稳分布? 假设 G 是连通的, 问经过所有顶点所需步数的期望值? 关于图上随机游动一个很好的综述见 Lovász [71]. 矩阵幂级数是矩阵分析 (matrix analysis) 这一课题的一部分. 有关矩阵幂级数更多的内容, 请见如 Horn and Johnson [58] 书中的第 5 章. 为避免用到矩阵范数理论, 书中定理 3.4 的证明看起来有一点“幼稚”.

第 3 章练习

1. 设 G 是一个 (有限) 图, 其顶点为 v_1, \dots, v_p . 假设由 (3.1) 定义的概率矩阵 $M(G)$ 的某一次幂的所有元都为正数. (不难发现这等价于 G 是连通的且包含至少一个奇数长度的圈, 但读者不需要证明这一结论.) 用 d_k 表示顶点 v_k 的度 (关联的边数). 令 $D = d_1 + d_2 + \dots + d_p = 2q - r$, 其中 G 有 q 条边且有 r 个自环. 从 G 的任一顶点出发, 在 G 的顶点上进行如本章定义的随机游动. 用 $p_k(\ell)$ 表示 ℓ 步后到达顶点 v_k 的概率. 假定 Perron-Frobenius 定理 (定理 3.3) 成立, 证明

$$\lim_{\ell \rightarrow \infty} p_k(\ell) = d_k/D.$$

称上面这一在 G 的顶点集上的极限概率分布为随机游动的平稳分布 (stationary distribution).

2. (a) 设 G 为有限图 (允许自环和重边). 假设存在某一整数 $\ell > 0$, 使得从任一给定顶点 u 到任一给定顶点 v 长为 ℓ 的游动数与 u 和 v 无关. 证明 G 中任意两顶点间都有 k 条边 (包括每个顶点都有 k 个自环).
- (b) 设 G 为具有如下性质的有限图 (允许自环和重边). 存在某一整数 $\ell > 0$ 使得如果从 G 的任一顶点出发做 ℓ 步的随机游动 (如前文所指), 那么到达任一顶点的概率相等. 换句话说, 若 G 有 p 个顶点, 则游动到达任一顶点 v 的概率都是 $1/p$. 证明有如 (a) 一样的结论, 即 G 中任意两顶点间都有 k 条边.
3. (a) 设 $P(x)$ 为一个实系数的非零多项式. 证明下面两条件等价.
- 存在一个实系数的非零多项式 $Q(x)$ 使得 $P(x)Q(x)$ 的所有系数都非负.
 - 不存在实数 $a > 0$ 使得 $P(a) = 0$.
- (b) (困难) 设 G 为连通的 (connected) 有限图, M 为 (3.1) 定义的概率矩阵. 证明下面两条件等价.
- 存在 \mathbb{P} 上的一个概率分布 P (那么选取 $k \in \mathbb{P}$ 的概率为 $P(k)$), 使得若我们首先从分布 P 中选取 k , 然后从 G 中的任一顶点出发随机游动 k 步, 则到达 G 中任一顶点的概率相等.
 - 图 G 是正则的, 且 M 不存在除 1 之外的正实数特征值.
4. (*) 固定 $0 \leq p \leq 1$. 从 n 维立方体 C_n 的顶点 $(0, 0, \dots, 0)$ 出发, 根据下面的规则沿着立方体的边游动: 每一个单位时间后, 要么以概率 p 停在原地, 要么随机 (均匀地) 跳到一个相邻的顶点. 因此跳到一个特定相邻顶点的概率为 $(1-p)/n$. 问 ℓ 个单位时间后重新回到 $(0, 0, \dots, 0)$ 的概率 $P(\ell)$ 是多少? 例如, $P(0) = 1$ 和 $P(1) = p$. 用一个有限和式表示所求概率.
5. 这一问题与本书没有直接的关系, 但它是一个经典的问题, 并有着非常智慧和漂亮的

结论. 设 G 是顶点集为 \mathbb{Z}_n (模 n 的整数) 的图, 对所有的 $i \in \mathbb{Z}_n$ 在 i 和 $i+1$ 之间有一条边. 因此 G 正好是一个 n -圈. 从顶点 0 出发, 并进行本书所述的随机游动, 这样从顶点 i 游动到 $i-1$ 或 $i+1$ 的概率都为 $1/2$. 对每一个 $i \in \mathbb{Z}_n$, 问顶点 i 是最后一个首次经过的顶点的概率是多少? 即当首次到达顶点 i 时, 已经经过了所有其他的顶点至少一次. 例如, $p_0 = 0$ (若 $n > 1$), 因为顶点 0 是第一个 (first) 经过的顶点.

6. (a) 证明: 若 u 和 v 是连通图 G 的两个顶点, 则不一定有 $H(u, v) = H(v, u)$, 其中 H 表示到达时间. 如果 G 是正则的情况又会怎样?
- (b) (困难) 对每一个 $n \geq 1$ 和有 n 个顶点的连通简单图的两顶点 u 及 v , 问 $H(u, v) - H(v, u)$ 可取到的最大值是多少?
7. (*) 设 u 和 v 为完全图 K_n 中不同的顶点. 证明 $H(u, v) = n - 1$.
8. (*) 设 P_n 是顶点为 v_1, \dots, v_n 的图, 对所有的 $1 \leq i \leq n - 1$, 在 v_i 和 v_{i+1} 之间有一条边. 证明 $H(v_1, v_n) = (n - 1)^2$. 对任意的 $i \neq j$, 问 $H(v_i, v_j)$ 等于多少? 若在 v_1 和 v_n 之间也有一条边情形又会怎样?
9. 设 K_{mn} 为一个完全二部图, 其顶点分为了两个部分 (A_1, A_2) , 其中 $\#A_1 = m$ 和 $\#A_2 = n$. 计算每一对相异顶点间的到达时间 $H(u, v)$. 存在两种不等价的情形: u 和 v 都在同一个 A_i 中, 或它们在不同的 A_i 中.
10. (*) 对图 G 的任意三个顶点 u, v, w , 证明

$$H(u, v) + H(v, w) + H(w, u) = H(u, w) + H(w, v) + H(v, u).$$

11. 设 $k \geq 0$, 且 u 和 v 为图 G 的顶点. 定义到达时间的 k 次二项矩 (binomial moment) $H_k(u, v)$ 为 $\binom{n}{k}$ 的平均值 (期望), 其中 n 是从 u 出发的随机游动首次到达 v 时所走的步数. 因此, 在 (3.2) 的记号下, 我们有

$$H_k(u, v) = \sum_{n \geq 1} \binom{n}{k} p_n.$$

设 x 为一个不定元. 沿用 (3.4) 中的记号, 证明

$$\sum_{k \geq 0} H_k(u, v) x^k = ((I_{p-1} - (x+1)\mathbf{M}[v])^{-1}T[v])_u.$$

12. (*) 本题是练习 3.7 的推广, 证明对完全图 K_n 的任意两个不同的顶点 u 和 v , 到达时间的 k 次二项矩为 $H_k(u, v) = (n-1)(n-2)^{k-1}$, $k \geq 1$. (当 $n=2$ 且 $k=1$ 时, 应取 $0^0 = 1$.)