# 第五讲 线性代数基础与应用

樊潇彦

复旦大学经济学院

# 本讲主要内容

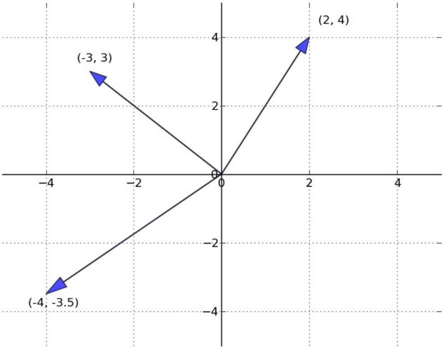

1. 向量、向量空间、内积和范数
   1.1 基本概念
   1.2 应用：范数和聚类
2. 矩阵
   2.1 基本概念
   2.2 应用：解方程组和投入-产出分析
3. 特征值和特征向量
   3.1 基本概念
   3.2 应用：变化中的稳态
4. 前沿应用：网络经济学
   4.1 无处不网络
   4.2 网络与矩阵

# 向量 $x$：点坐标

向量（vector）：$x \in \mathbb{R}^N$ 是 $N$ 维实数空间中的点坐标。

一般我们使用列向量：

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}$$

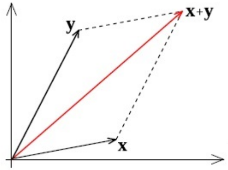


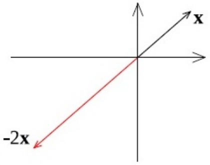

转置（transpose）后得到行向量：

$$x^T = (x_1, x_2, \ldots, x_N)$$

# 向量空间 $V$：定义了加法和纯量乘法的实空间

- 对向量 $x, y \in \mathbb{R}^N$ 和实数 $a \in \mathbb{R}$，定义向量的加法和纯量乘法：




$z = x+y$, $z_i = x_i+y_i$　　　$z = ax$, $z_i = ax_i$

- 向量空间（Vector space，又称线性空间 Linear space）是对向量加法和纯量乘法封闭的实空间的子集，即：
  - $V \subset \mathbb{R}^N$
  - 对 $\forall x, y \in V$, $a \in \mathbb{R}$，有 $x + y \in V$、$ax \in V$

# 内积和范数

- 对于 $x, y \in V$，定义内积（inner product，也称点积 dot product），例如总支出就是价格向量和消费向量的内积：

$$C = p \cdot c = p^T c = \sum_{i=1}^{N} p_i c_i$$

- 范数（norm）：向量的长度（length），或点到原点的距离（distance）。
  - 欧氏范数（Euclidean norm）：

$$\|x\| = \sqrt{x \cdot x} = \left( \sum_{i=1}^{n} x_i^2 \right)^{1/2}$$

  - $p$ 范数（$p$-norm）：

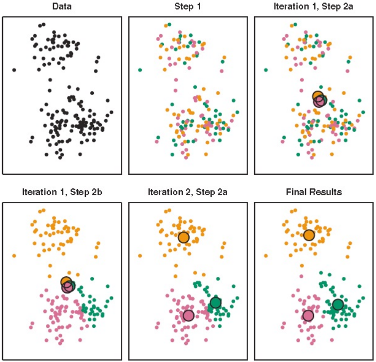
$$\|x\|_p = \left( \sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

# 基于欧氏距离的聚类分析

- 聚类分析（clustering）是在一个数据集中寻找子群或类的技术。
- K均值聚类（K-means Clustering）是将观测划分到事先规定的 $K$ 个类中，使类内差异尽可能小。公式为：

$$\min \sum_{k=1}^{K} W\left(C_k\right) = \frac{1}{|C_k|} \sum_{u,v \in C_k} \|x_u - x_v\|$$

其中 $\|\cdot\|$ 为欧几里得距离，$|C_k|$ 为第 $k$ 类所包含的样本数。




Source: G. James, D. Vitten, T. Hastie, R. Tibshirani(2015)

# 矩阵：定义

$N \times K$ 维矩阵 $X$ 记为：

$$X_{N\times K} = \begin{bmatrix} x_{11} & \cdots & x_{1K} \\ \vdots & \ddots & \vdots \\ x_{N1} & \cdots & x_{NK} \end{bmatrix}$$

对于一个有 $N$ 个样本、$K$ 个变量的数据集 $X$，我们可以将其视为 $\mathbb{R}^K$ 维空间中的 $N$ 个点，每个行向量都是点的坐标；或者 $\mathbb{R}^N$ 维空间中的 $K$ 个点，每个列向量为点的坐标。

$$X_{N\times K} = \begin{bmatrix} \tilde{X}_1 \\ \vdots \\ \tilde{X}_N \end{bmatrix} = [X_1 \ldots X_K]$$

# 矩阵与向量的乘法

- 矩阵 $X_{N\times K}$ 右乘列向量 $\beta_{K\times 1}$，得到的列向量 $y_{N\times 1}$ 是 $X$ 的列向量的线性组合：

$$y = X\beta = \sum_{k=1}^{K} X_k \beta_k$$

- 行向量 $\tilde{\alpha}_{1\times N}$ 左乘矩阵 $X_{N\times K}$，得到的行向量 $z_{1\times K}$ 是 $X$ 的行向量的线性组合：

$$z = \tilde{\alpha} X = \sum_{i=1}^{N} \alpha_i \tilde{X}_i$$

因此 $X_{N\times K}$ 也被称为不同维度的线性空间之间的转换矩阵：

$$X\beta: \ \mathbb{R}^K \to \mathbb{R}^N$$
$$\tilde{\alpha} X: \ \mathbb{R}^N \to \mathbb{R}^K$$

# 特殊矩阵

- 单位阵：

$$I=\begin{bmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix}$$

- 对称方阵：

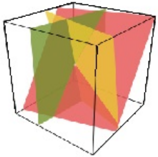
$$A=A^T \quad \Leftrightarrow \quad a_{ij}=a_{ji}$$

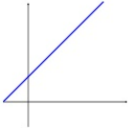
- 可逆阵：如果对于方阵 $A$ 和 $B$ 下式成立，则 $A$ 和 $B$ 互为逆矩阵。

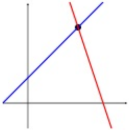
$$AB=BA=I \quad \Leftrightarrow B=A^{-1},\ A=B^{-1}$$

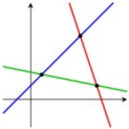
# 求解线性方程组：含义与定理

- 求解线性方程组 $Ax = b$：
  - 将 $b$ 表示为矩阵 $A_{N\times K}$ 的 $K$ 个列向量的线性组合；
  - 寻找矩阵 $A_{N\times K}$ 的 $N$ 个行向量所表示的超平面的交点 $x$：






- 解的判定定理：
  - $\text{rank}\,(A) < K$ 时有无穷多解（under-determined）；
  - $\text{rank}\,(A) = K = N$ 时，有唯一解（just-determined），$x = A^{-1}b$；
  - $\text{rank}\,(A) = K < N$ 时无解（over-determined），$A$ 为方阵时不可逆。

# 求解线性方程组：剑桥食谱

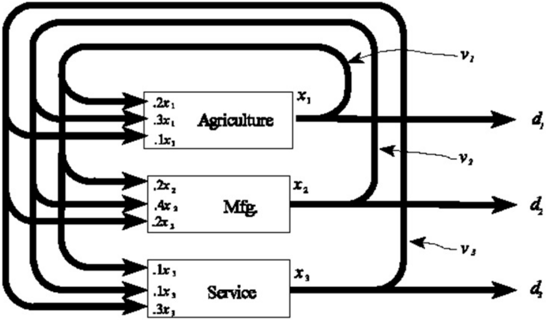
假定牛奶等三种食品每100克所含的营养成分，以及每天应摄入的总营养成分如下表所示：

| | 脱脂牛奶 | 大豆粉 | 乳清 | 应摄入 |
|---|---|---|---|---|
| 蛋白质 | 36 | 51 | 13 | 33 |
| 碳水化合物 | 52 | 34 | 74 | 45 |
| 脂肪 | 0 | 7 | 1.1 | 3 |

请制定一份健康食谱，列出三种食品的占比。

# 投入-产出表：经济背景

20世纪30年代美国经济学家瓦西里·列昂惕夫（W.Leontief）在前人关于经济活动相互依存的研究基础上提出了投入-产出表（Input-Output Table）的分析框架，1936年发表了“美国经济制度中投入产出数量关系”一文，1953 年与他人合作出版了《美国经济结构研究》一书，1973年获得诺贝尔经济学奖。



# 投入-产出表（Input-Output Table）：数学表述

| | | 中间需求 | | | | 最终需求 | 总产出 |
|---|---|---|---|---|---|---|---|
| | | 部门1 | 部门2 | ... | 部门N | 消费 投资 净出口 | |
| 中间投入 | 部门1 | $a_{11}x_1$ | $a_{12}x_2$ | $\cdots$ | $a_{1N}x_N$ | $d_1$ | $x_1$ |
| | 部门2 | $a_{21}x_1$ | $a_{22}x_2$ | $\cdots$ | $a_{2N}x_N$ | $d_2$ | $x_2$ |
| | ... | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| | 部门N | $a_{N1}x_1$ | $a_{N2}x_2$ | $\cdots$ | $a_{NN}x_N$ | $d_N$ | $x_N$ |
| 最终投入 | 工资 折旧 利润 税收 | $v_1$ | $v_2$ | $\cdots$ | $v_N$ | | |
| 总投入 | | $x_1$ | $x_2$ | $\cdots$ | $x_N$ | | |

- $x_{ij}$：第 $i$ 个部门对第 $j$ 个部门的中间投入
- $A$：直接消耗矩阵，元素 $a_{ij} = \frac{x_{ij}}{x_j}$
- $d_i$：第 $i$ 个部门的最终需求，等于支出法核算的GDP
- $v_j$：第 $j$ 个部门的最终投入，等于收入法核算的GDP
- $x_i$：第 $i$ 个部门的总产出（总需求），等于总投入（总供给）

# 投入-产出表：求解公式

- 中间需求+最终需求=总产出：

$$Ax + d = x \Leftrightarrow x = (I - A)^{-1} d$$

- 中间投入+最终投入=总投入：

$$Dx + v = x \Leftrightarrow x = (I - D)^{-1} v$$

$$D = \begin{bmatrix} \sum_{i=1}^{N} a_{i1} & 0 & \cdots & 0 \\ 0 & \sum_{i=1}^{N} a_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{i=1}^{N} a_{iN} \end{bmatrix}$$

# 投入-产出表：分析示例

| | 第一产业 | 第二产业 | 第三产业 | 最终需求 | 总产出 |
|---|---|---|---|---|---|
| 第一产业 | 15 | 20 | 30 | 35 | 100 |
| 第二产业 | 30 | 10 | 45 | 115 | 200 |
| 第三产业 | 20 | 60 | 0 | 70 | 150 |

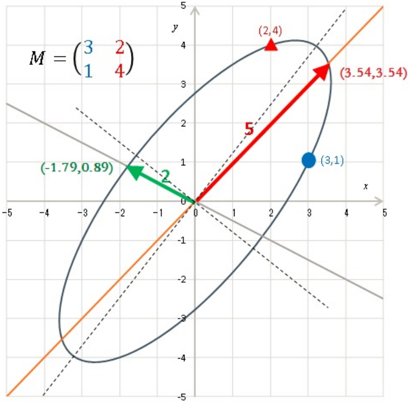
- 检验最终需求和总产出之间的关系公式；
- 假定投入产出关系不变，当最终需求变为 $(100, 200, 300)$ 个单位时，总产出等于多少？

# 特征值和特征向量

对于 $N$ 维方阵 $A$，如果存在常数 $\lambda$ 和非零向量 $x$，使得 $Ax = \lambda x$（或 $\tilde{x}A = \lambda\tilde{x}$），则称 $\lambda$ 是 $A$ 的特征值（eigen value），$x$（或 $\tilde{x}$）是对应于 $\lambda$ 的右（或左）特征向量（eigen vector）。

以二维方阵 $M$ 为例：

- 特征值是方阵的点所在的椭圆的长轴和短轴；
- 特征向量是经过其线性变换后，方向不变、长度伸缩的向量。



# 概率转移矩阵和平稳分布

- 如果随机变量 $z_t$ 从 $t$ 时刻的状态 $Z_i$ 变到 $t+1$ 时刻的状态 $Z_j$ 的概率与时间无关，则可以用概率转移矩阵 $P=[p_{ij}]$ 刻画它的变化规律，其中元素：

$$p_{ij}=P\left(z_{t+1}=Z_j|z_t=Z_i\right)$$
$$p_{ij}\geq 0,\quad \sum_{j=1}^{n}p_{ij}=1$$

- 对于定义在 $\mathbb{Z}$ 上的概率分布 $\pi=(\pi_1,\pi_2...\pi_N)$，如果 $\pi=\pi P$，则 $\pi$ 为 $z_t$ 的平稳分布。显然，$\pi$ 是 $P$ 的特征值1所对应的左特征向量。
- 从某一初始分布 $\pi^0=\left(\pi_1^0,\pi_2^0...\pi_N^0\right)$ 出发，有 $\pi^{t+1}=\pi^tP=\pi^0P^t$。如果 $\lim_{t\to\infty}\pi^t=\pi^*$ 存在，则称 $\pi^*$ 为 $z_t$ 的极限分布。如果 $\pi^*$ 与 $\pi^0$ 无关，则称该马氏链是遍历的（ergodic）。如果 $\forall i,j,\ p_{ij}>0$，则 $z_t$ 遍历平稳，且 $\pi^*=\tilde{P}_i^*$：

$$\lim_{t\to\infty}P^t=P^*=\begin{bmatrix}\pi_1 & \pi_2 & ... & \pi_N\\ ... & ... & ... & ...\\ \pi_1 & \pi_2 & ... & \pi_N\end{bmatrix}$$

# 稳态人口分布


David C. Lay著：《线性代数及其应用（原书第3版）》，刘深泉等译，机械工业出版社，2005

- 根据上图写出人口迁移的转移矩阵。
- 是否存在人口分布的稳态？
- 如果预计未来总人口为1000万，请问居住在城市和农村的各占多少？

# 一元一阶差分方程的解与平稳性条件

$y_t$ 的动态过程可以表述为一元一阶差分方程（其中 $\varepsilon_t \sim N\left(0,\sigma^2\right)$，$y_0=\mu+\varepsilon_0$）：

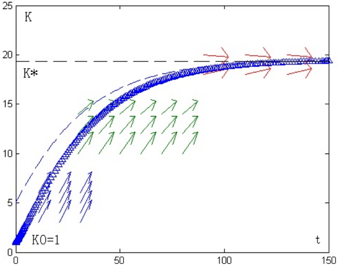
$$y_t=\mu+\phi y_{t-1}+\varepsilon_t=\mu\sum_{s=0}^{t}\phi^s+\sum_{s=0}^{t}\phi^s\varepsilon_{t-s}$$

由于 $\frac{\partial y_t}{\partial\varepsilon_0}=\phi^t$，当时间 $t\to\infty$ 时，$y_t$ 的平稳性意味着当期的随机冲击影响有限，这等价于 $|\phi|<1$。否则 $\phi>1$ 时 $y_t$ 会单调发散，$\phi<-1$ 时会震荡发散。

例如在 Solow-Swan 模型中，一个国家的资本积累可以表述为非线性一元一阶差分方程：

$$K_{t+1}=(1-\delta)K_t+sY_t=(1-\delta)K_t+sAK_t^{\alpha}$$

给定参数 $\delta=0.05$，$s=0.4$，$A=1$，$\alpha=0.3$，右图展示了 $K_0=1$ 和 $K_0=5$ 两种情况下资本 $K_t$ 的动态演进过程。



# 多元一阶差分方程组的解与平稳性条件

一元高阶差分方程 $y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} ... + \phi_p y_{t-p} + \varepsilon_t$ 可以用降阶增元的方法改写为多元一阶差分方程组 $Y_t = \tilde{\mu} + \Phi Y_{t-1} + \tilde{\varepsilon}_t$：

$$\begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix} = \begin{bmatrix} \mu \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_p \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

上述方程组也可以写为 $(I - \Phi L) Y_t = \tilde{\mu} + \tilde{\varepsilon}_t$，根据汉密尔顿(1999)，当 $\Phi$ 的所有特征根的模都小于1（$|\lambda_i| < 1$，$(i = 1, 2 ... p)$）时，方程组的解为：

$$Y_t = (I - \Phi L)^{-1} (\tilde{\mu} + \tilde{\varepsilon}_t)$$

其中 $(I - \Phi L)^{-1} = I + \Phi L + (\Phi L)^2 + ... = \sum_{j=0}^{+\infty} (\Phi L)^j$。

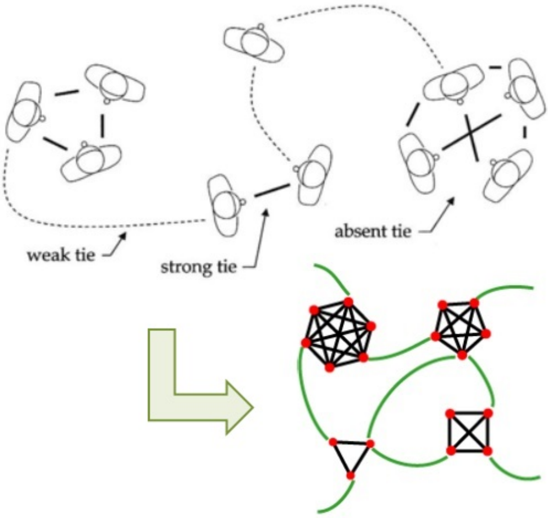
# 两部门宏观经济模型

假定一个国家的宏观经济可以表述为下面的差分方程组：

$$
\begin{aligned}
Y_t &= C_t + I_t \\
C_t &= C_0 + cY_{t-1} \\
I_t &= I_0 + \beta\left(C_t - C_{t-1}\right)
\end{aligned}
$$

- 请将上述模型改写为消费和投资的二元一阶差分方程组；
- 如果 $c = 0.9$，$\beta = 0.5$，该国的宏观经济是否稳定？
- 假如 $C_0 = 100$，$I_0 = 1000$，请问该国长期的稳态产出、消费和投资分别为多少？

# 社会学中的关系网络

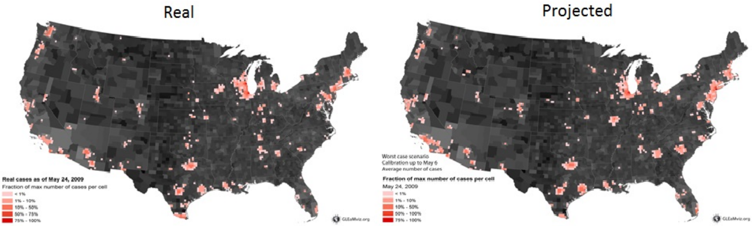



图片来源：Barabasi(2013)。

# 超越传统经济学[1]

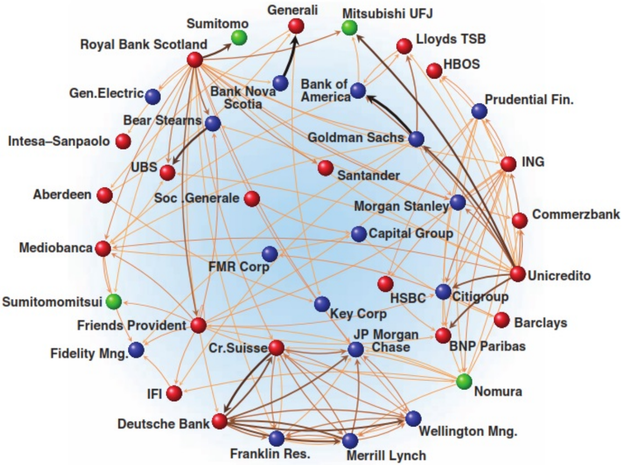
- 传统经济学认为：
  - 行为人是匿名的（anonymous agents）的；
  - 行为人在集中市场（centralized markets）中交易。
- 而事实上：
  - 交易者的身份（identity）很重要；
  - 交易者之间的互动（interaction）对于信息传递和交易结果有重要影响。
- 传统经济学所忽略的正是网络经济学研究的重点。

# 网络的力量：预测H1N1的传播

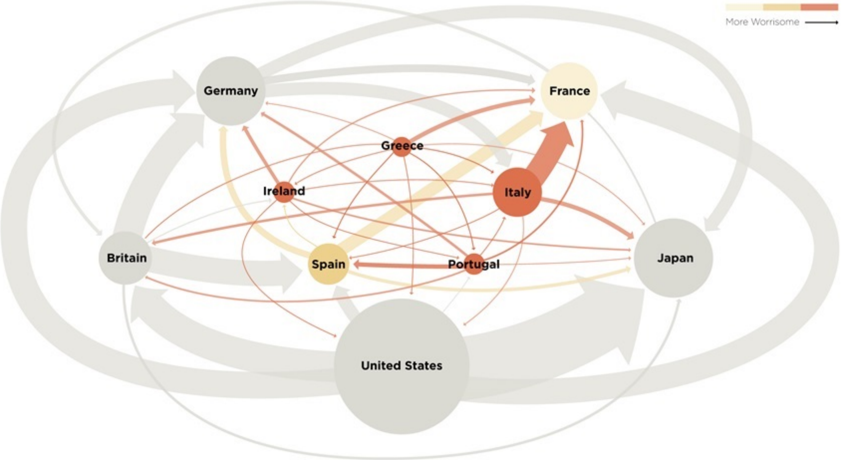

注：Barabasi(2013)。

# 网络的力量：银行间融资网络与金融危机




图片来源：Schweitzer et al.(2009, Science)。

# 网络的力量：国家间借贷网络与欧债危机

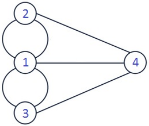



图片来源：Barabasi(2013)。

# 网络的数学表示：图和矩阵

- 哥尼斯堡七桥问题：

- 用矩阵描述图：



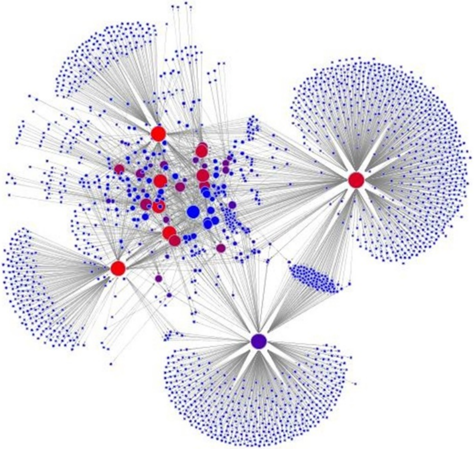

$$N=\{1,2,3,4\} \quad g=\begin{bmatrix} 0 & 2 & 2 & 1 \\ 2 & 0 & 0 & 1 \\ 2 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

- 网络（network）和图（graph）$G=(\mathbf{N},g)$ 的基本要素：
  - 节点（nodes）：集合$\mathbf{N}=\{1,2,...N\}$，指行为人或组织。
  - 连接（links）：邻接矩阵 $g=[g_{ij}]_{i,j\in N}$ ，表示指节点之间的关系。

# 常用统计指标：度

- 度（degree，$d_i$）：与节点相关联的边数，有向图中分为入度（in-degree，$d_j^{in} = \sum\limits_{i=1}^{N} g_{ij}$）和出度（out-degree，$d_i^{out} = \sum\limits_{j=1}^{N} g_{ij}$）；
- 平均度（average degree），$\langle d \rangle = \frac{1}{N}\sum_{i=1}^{N} d_i$：对于无向图$\langle d \rangle = \frac{2L}{N}$，对于有向图$\left\langle d^{in} \right\rangle = \left\langle d^{out} \right\rangle = \frac{L}{N}$。

# 真实网络：明星和枢纽



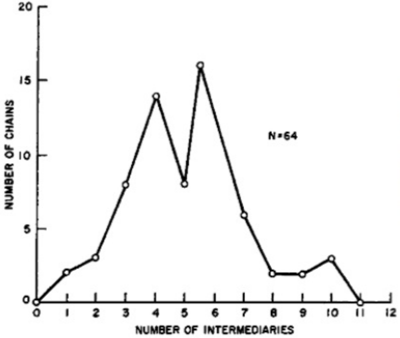
图片来源：Adamic(2013)课件。

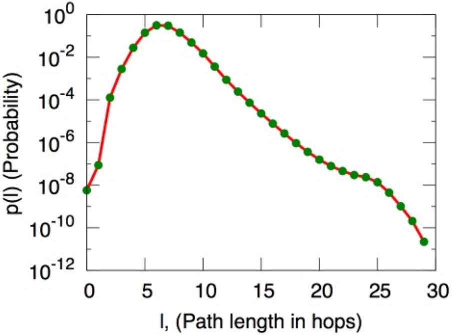
# 常用统计指标：直径与距离

- 直径（diameter，$l_{max} = \max\{l_{ij}\}$）：网络中任意两个节点之间的最大距离。不连通的网络直径无穷大。
- 平均距离（average distance，$\langle l \rangle = \frac{2\sum_{j \neq i} l_{ij}}{N(N-1)}$），对于不连通的网络，可以定义为各个连通分量平均距离的加权平均。

# 真实网络：六度分隔

电影《Six Degrees of Separation》：不管是美国总统还是威尼斯的船夫，只要找到正确的5个人，我们就能联系起来。




注：左图为Travers and Milgram(1969)试验中成功将信送达目标人物的路线长度分布图，右图为微软MSN全球活跃用户相隔距离分布图，两图均来自Easley and Kleinberg(2010)。

# 常用统计指标：中心性

- 度中心性（degree centrality）：节点在网络中的连接程度，如网络名人

$$Ce_i^d = \frac{d_i}{N-1}$$

- 亲近中心性（closeness centrality）：节点抵达其他节点的难易度，如八卦传播者

$$Ce_i^c = \frac{N-1}{\sum_{j\neq i} l_{ij}}$$

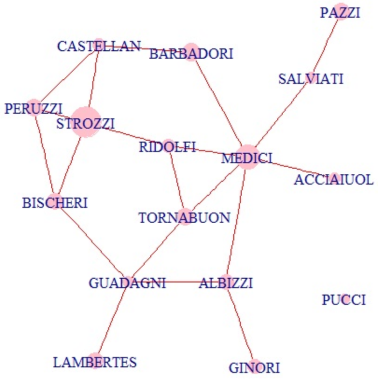
- 中介中心性（betweenness centrality）：节点在连接其他节点方面的重要性，如跨界者和社群桥梁（其中$P(kj)$是$k$和$j$之间最短路径的个数，$P_i(kj)$ 是$i$位于多少条这样的最短路径上）：

$$Ce_i^b = \sum_{k\neq j, i\notin(k,j)} \frac{P_i(kj)/P(kj)}{(N-1)(N-2)/2}$$

- 特征向量中心性（eigenvector centrality）：节点有多少重要的邻居，如灰衣主教和幕后大佬（其中$\lambda$ 是邻接矩阵$g$ 最大的特征值）：

# 真实网络：美第奇家族的崛起

回顾15世纪佛罗伦萨家族的政治地位问题，从财富总量来看Medici不如Strozzi（图中节点的大小表示财富水平），但是美第奇家族为什么最终会崛起呢？




注：15世纪佛罗伦萨的婚姻网络图，根据Wasserman and Faust(1994)数据，用Pajek绘制。

# 真实网络：GOOGLE的法宝

Wiki: The PageRank values are the entries of the dominant right eigenvector of the modified adjacency matrix.