

Problem Set #1

Huanye Liu

- (a) The data can be obtained by visiting the website <http://nber.org/patents/> and clicking the download link `apat63_99.zip` from the table at the bottom of the page. The original source of data is the collection of patent records from USPTO and the following paper:

Hall, B.H., A. B. Jaffe, and M. Trajtenberg(2001). "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools." NBER Working Paper 8498

The curator is NBER, and the PI for this project is Iain Cockburn..

- (b) Other key papers that have used this data include:

Hall, Bronwyn H., Adam Jaffe, and Manuel Trajtenberg. "Market value and patent citations." *RAND Journal of economics* (2005): 16-38.

Acs, Zoltan J., Luc Anselin, and Attila Varga. "Patents and innovation counts as measures of regional production of new knowledge." *Research policy* 31.7 (2002): 1069-1085.

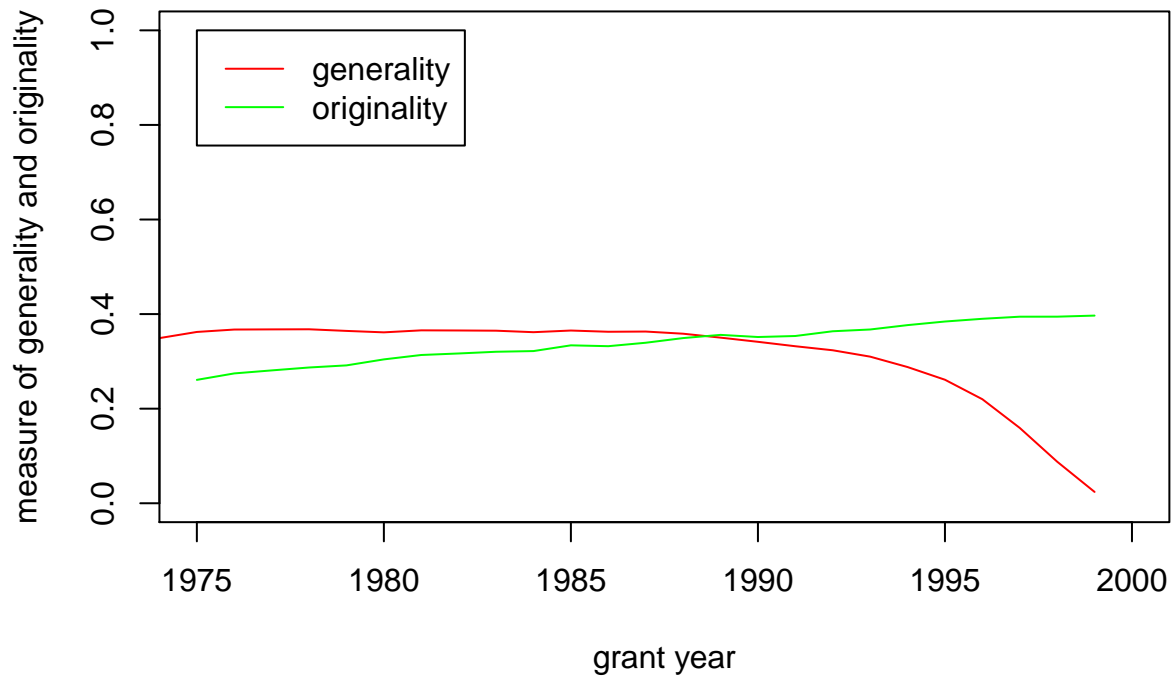
Leskovec, Jure, Jon Kleinberg, and Christos Faloutsos. "Graphs over time: densification laws, shrinking diameters and possible explanations." *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*. ACM, 2005.

- (c) USPTO makes record of the 10 original variables of patent information based on the USPTO's TAF database granted during the period 1963 to December 1999. Hall, Bronwyn H., Adam Jaffe, and Manuel Trajtenberg make records of 13 new variables of patent information.

- (d) The table of the descriptive statistics is shown as below:

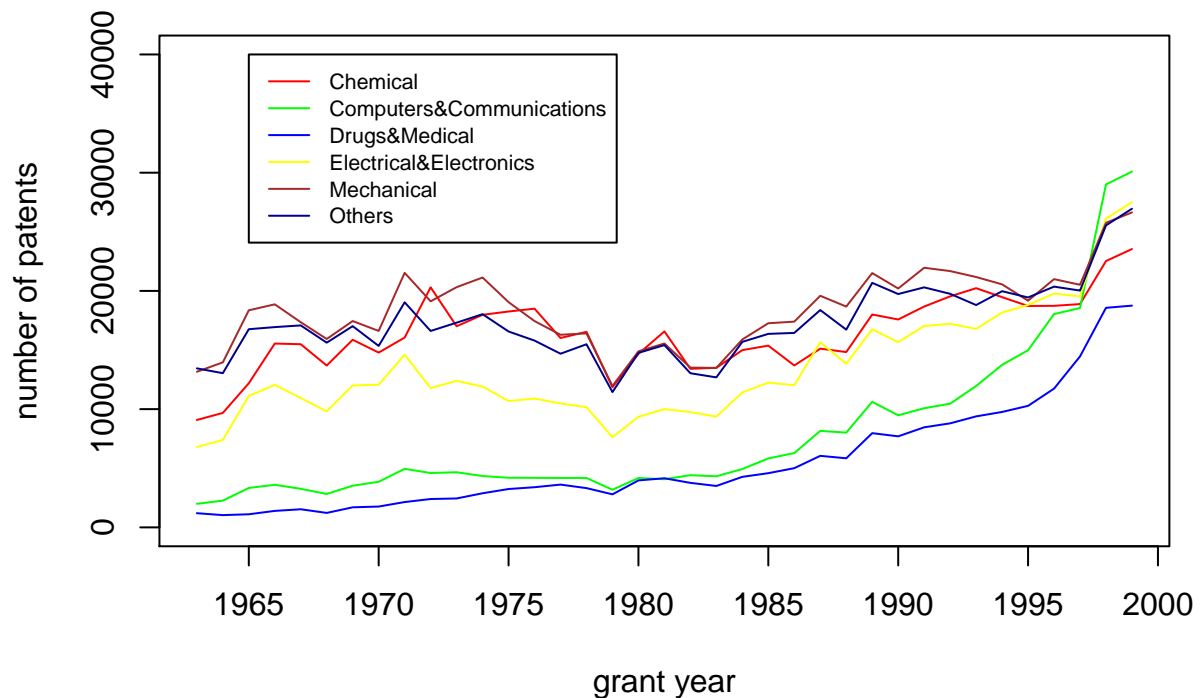
Variables	Mean	Std	Median	Min	Max
Number of Citations Made	7.720	9.000	6.0	0	770
Number of Citations Received	4.779	7.346	3.0	0	779
Mean Forward Citation Lag	8.306	5.804	7.0	0	96
Mean Backward Citation Lag	14.100	11.769	10.5	0	154
Share of Self-Citations Made - Upper Bound	0.136	0.256	0.0	0	1
Share of Self-Citations Made - Lower Bound	0.110	0.218	0.0	0	1
Share of Self-Citations Received - Upper Bound	0.132	0.260	0.0	0	1
Share of Self-Citations Received - Lower Bound	0.125	0.250	0.0	0	1

Patent originality and generality time series



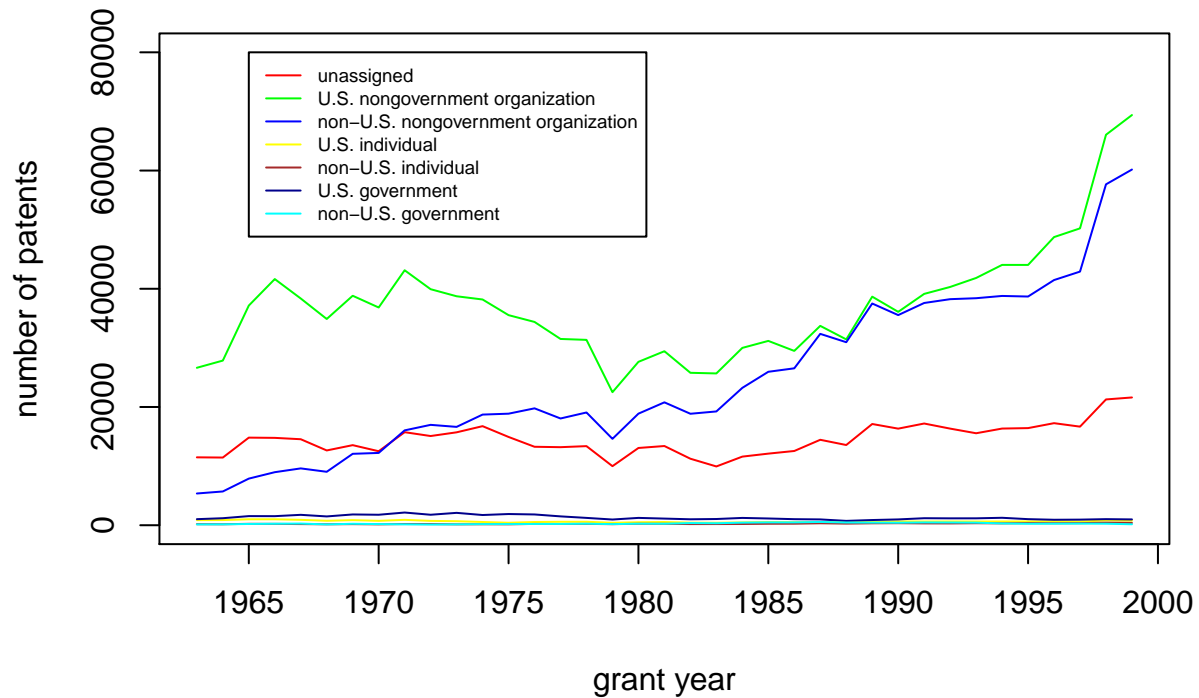
- (e) As the above graph shows, we can observe the decreasing trend of patent generality and increasing trend of patent originality during the period 1975 to 1999. The former trend may be explained by the fact that research becomes more focused on some specific domain to attain some specific goal in recent years, and the latter trend may be due to all kinds of incentives and rewards to encourage research with high level of originality in recent years.

Number of patents time series in six technological categories



- (f) As the graph shows above, we can see a general increasing number of patents in all six technological categories. However, compared with other categories, the increasing trend is more consistent in category Computers&Communications and category Drugs&Medical, and the number of patents in all six categories climbs up sharply from year 1995 to 2000.

Number of patents time series of different assignee types



- (g) We can see from the above graph that the number of patents assigned to the nongovernment organization is way bigger than those assigned to its government and individual counterparts. In addition, both numbers of U.S. and non-U.S. nongovernment patent assignments increase dramatically since 1985, while patent assignments to the individual and to the government keep at a much lower level with no obvious increase.