Why Help Those Weaker Than You?

Tomas Ukkonen, Novel Insight Research

tomas.ukkonen@novelinsight.fi

Abstract

In recent years, the ideas of the political right, where, for example, selfishness is good for society and that there should not be many taxes to have social safety nets, have gained more ground. However, there are mathematical results that show that thinking, in which only the strongest should survive, may be partly wrong. In this article, we describe game theoretic, global optimization, bayesian and information-theoretic, and risk based mathematical results that infer it is often optimal to help others.

1 Global Optimization

Optimization[1], which often reduces to a simple competition-based optimization (gradient ascend), gets stuck to **local maximums** (See Figure 1). This means the good optimization of individuals in groups may require more than relatively simple changes, the competition, and the survival of the fittest. To escape from a local maximum, a search through potentially worse solutions is required. This means that the weaker ones (initially potentially better solutions) should be supported by the stronger ones. (If we add money to this, individuals can earn money at the local maximum and then do an expensive search of new better optimum using extra resources they have. But in practice, it may be difficult for individuals to change, and we currently cannot change our genes or brains easily but only give birth to new children with altered traits and genes.)

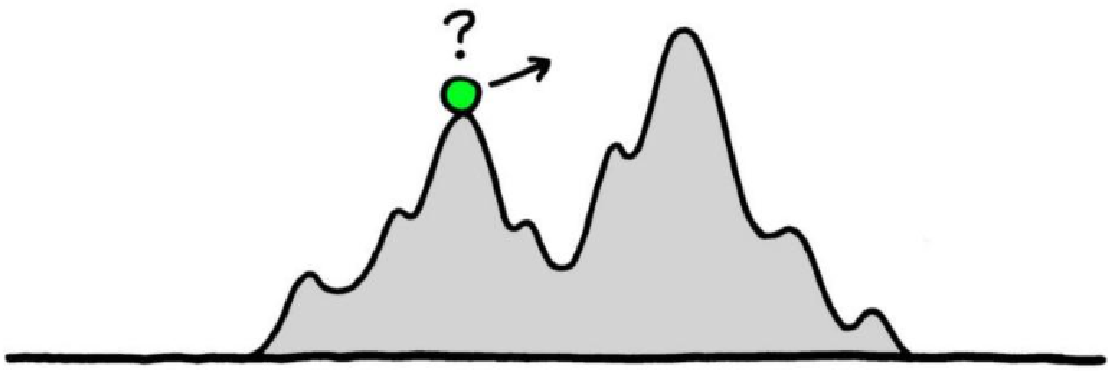

Figure 1.

2 Game Theory

In **game theory**[2], there is a concept of **the price of anarchy** in which a group of selfish agents does worse than a centrally controlled optimal system. While the optimal central control, which doesn't waste resources to **the internal competition**, may be a bit unrealistic, with the increasing capacity of computers and better and more accurate databases, we might eventually have this optimal system.

Another game-theoretic result is that in competition, the best solution is not a single best type of individual but **a probability distribution** of different choices/individuals (**mixed strategies, hawk-dove example,** (you don't need the nash equilibrium and just need to optimize for the best-expected value)). A society, in which the whole population is CEOs, cannot function, and instead, we need different kinds of people: workers, entrepreneurs, scientists, businessmen, engineers, etc. Therefore, the stronger should support the weaker ones through a complicated system of laws and regulations.

3 Optimal Adaptation Speed

Additionally, if an environment changes and the population must adapt as fast as possible, **information theory**[3] shows that the fastest way to adapt (in the information-theoretic sense [has few problems, tails don't carry very much information for example]) to a new situation, is to calculate using **bayesian inference**[4]. This means calculations/optimization using population probability distributions again and the tails of the less probable individuals/genes must be supported somehow.

Bayesian inference can be shown to be information theoretically optimal.

$$p(x||\text{data})p(\text{data}) = p(\text{data}||x)p(x)$$

$$\log(p(x||\text{data})) + \log(p(\text{data})) = \log(p(\text{data}||x)) + \log(p(x))$$

$$H(X||\text{DATA}) + H(\text{DATA}) = H(\text{DATA}||X) + H(X)$$

$$H(X||\text{DATA}) = H(X) - I(X;\text{DATA})$$

However, in practice information theory is no good measure when comparing distributions using Kullback-Leibler divergence (more general case of mutual information). Another, better way to compare distributions is by taking the maximum ratio of different probability values.

$$D'_{\text{KL}} = \sum p_i \log\left(\max\left(\frac{p_i}{q_i}, \frac{q_i}{p_i}\right)\right) = \sum p_i \left|\log\left(\frac{p_i}{q_i}\right)\right|$$

4 Risk

Yet another result according which we should help each other is to make **risk-taking** less dangerous. In a society, in which we can take rational risks (increase standard deviation of the outcome σ), the expected returns (mean μ) of the actions increase too and society as a whole progresses faster. Groups can increase this variance/risk, for example, by forming larger organisations like governments and companies, which have more resources (money against **risk of ruin**). They can then take larger risks than individuals themselves and have much larger expected returns.

It is also possible to protect from risks and disasters (significantly reducing the variance of the outcomes) by taking insurance. **Insurance providers therefore genuinely help people.** Insurance companies and insurance takers both benefit from this relationship. This relationship exists often when the stronger or larger entity such as a company or government helps individual persons which pay small fees or taxes for the help. For more details read Section 6 about *Insurance Statistics*.

5 Evolution and Competion Fails

The current theory of evolution tells that biological organisms have evolved through competion and survival of the fittest. But in practice, genetic algorithms[5] in computers cannot adapt or optimize to a large number of genes. This means that competion can at most find well-adapted organisms having less than 30 binary genes. In normal organisms, there are 1000 or much more genes, so evolution cannot produce real-life organisms. This is because optimizing for 2^{1000} combinations of gene values is impossible using simple evolution/competion. Similar observations can be seen in companies and market economies although in economics we have humans with high-quality optimization capabilities for their companies.

One practical example of the capitalist system failure is **the USA's healthcare**, which is one of the most expensive in the World, even when competion in USA should in theory reduce the costs down.

Bibliography

1. A First Course in Optimization Theory. Rangarajan K. Sundaram. 1996. Cambridge University Press.

2. Learn Game Theory: A Primer to Strategic Thinking and Advanced Decision-Making. Albert Rutherford. 2021. ARB Publications.
3. Information Theory: A Tutorial Introduction. James Stone. 2015. Sebtel Press.
4. Bayesian Data Analysis. Andrew Gelman et al. 2013. Chapman and Hall/CRC.
5. Introduction to Genetic Algorithms. S.N. Sivanandam and S.N. Deepa. 2007. Springer.

6 Insurance Statistics

In insurance mathematics, insurance takers protect against disasters that cost B with probability p and have a pleasant outcome A with probability $1-p$. The cost of insurance is C and it pays disaster costs B fully if a disaster happens. Now let the pleasant outcome $A=0$ and this means the expected return R without and with insurance is:

$$E[R||\text{insurance}=\text{no}]=(1-p)A-p\,B=-p\,B$$
$$E[R||\text{insurance}=\text{yes}]=(1-p)A-C=-C$$

The profit of insurance S per insurance taker for the company providing them is

$$S=C-p\,B$$

The insurance provider wants a $q\%$ return for each time interval from its insurance S profit.

$$q=\frac{C-p\,B}{p\,B}$$

Now assume disaster cost is $B=10^5$ and its probability is $p=10^{-5}$ (house is destroyed) so

$$E[R||\text{insurance}=\text{no}]=-p\,B=\text{-1}$$

This means insurance cost must be ($q=1.05$ (5% profit per time interval)):

$$C=(q+1)\,p\,B=2.05$$

Therefore profit for the insurance is $S=1.05$ and $E[R||\text{insurance}=\text{yes}]=-2.05$

Insurance is a profitable business for the insurance giver but the customer's expected returns of investment are smaller if the insurance is taken (-1 vs -2.05). This means that if the insurance taker has an infinite amount of money (s)he doesn't need to take the insurance.

In practice C has no variance and has no risk but if no insurance is not taken, there is the risk of disaster. We assume R is normally distributed and calculated standard deviation $\text{StDev}[R||\text{no}]=\sqrt{E[R^2]-E[R]^2}=\sqrt{p\,B^2-(-1)^2}\approx 10^{2.5}=316,2$. If the insurance is not taken the worst case scenario (two standard deviations below the mean outcome of normal distribution means 97.5% probability of worst case) is

$$E[R||\text{insurance}=\text{no}]-2\,\text{StDev}[R||\text{insurance}=\text{no}]=-1-2*316.2=-633,5$$
$$E[R||\text{insurance}=\text{yes}]-2\,\text{StDev}[R||\text{insurance}=\text{yes}]=-2.05-2*0=-2.05$$

This means that by taking insurance we can reduce risks significantly and have 631,5 better worst-case expenses. If the money we have per time interval is less than 633,5, we can take the insurance and have the worst-case expense that is only -2.05 which means by taking insurance we can survive disasters.

On the other hand, the disaster provider gives insurance to a large number of people meaning that the profit per customer $S = C - p\,B$ is close to the expected value and is approximately normally distributed ($S = \frac{1}{N}\sum_{i=0}^{N} S_i,\ E[S] = C - p\,E[B],\ \mathrm{Var}[S] = \frac{1}{N}\mathrm{Var}[p\,B] = \frac{1}{N}\,p\,B^2$). If $N = 10^6$, then $\mathrm{StDev}[S] = 0.316$ and $E[S] = 1.05$. The average profit from a customer is between [0,417;1.632]. This means that if we have a million customers then the average profit per time interval is very likely to be positive. **Insurance giver improves society by helping insurance takers survive from disasters, hazards, and risks**.

The average amount of money required per time interval is $N\,p\,B = 10^6$. This means the insurance company must have 1 million euros or a bit more to handle peaks in disasters.