Why Min, Max, Opening, and Closing Stock Prices Are Empirically Most Appropriate for Predictions, and Why Their Linear Combination Provides the Best Estimate for Beta

Somsak Chanaim[1], Olga Kosheleva[2], and Vladik Kreinovich[2]

[1]International College of Digital Innovation, Chiang Mai University
Chiang Mai 50200, Thailand, somsak ch@cmu.ac.th
[2]University of Texas at El Paso, El Paso, Texas 79968, USA
olgak@utep.edu, vladik@utep.edu

1. Machine learning – currently the best way to predict stock prices

- We want to select the best investment strategy.
- So, it is important to predict future prices of different financial instruments.
- In the past, complex analytical models were used to predict future stock prices.
- However, these models, whether they are linear or nonlinear, provide only an approximate description of the corresponding dynamics.
- The real dynamics is much more complex.
- It is therefore reasonable to use prediction techniques which are not limited to any specific class of models.
- Such techniques are known as *machine learning* techniques.
- At present, machine learning techniques – usually, techniques of *deep learning* – provide the best way to predict stock prices.

2. What input should we use for prediction?

- Traditionally, most financial markets report closing daily prices of different financial instruments.
- Sometimes, opening prices are also reported.
- At present, with everything online, one can trace moment-by-moment changes in the price of each instrument.
- At first glance, it may seem that:
  - the more information we use,
  - the more accurate the predictions will be.
- To some degree, this is true.
- If we start with scarce data and add more data, we get more and more accurate predictions.
- However, after a while, adding more data becomes counter-productive, for two reasons.

3. What input should we use for prediction (cont-d)

- First, data comes with noise: for example:
  - a significant part of moment-by-moment fluctuations in prices is caused
  - by short-term traders trying to benefit from small changes in prices.
- These changes do not help in predicting longer-term trends.
- They only obscure the picture.

4. What input should we use for prediction (cont-d)

- Second, by their structure, deep neural networks cannot input too much data.
- If you try to feed too much data:
  - they will compress it anyway,
  - by using general data compression techniques.
- From this viewpoint, it is definitely better to perform compression tailored to the application area.
- This will lead to the smallest possible information loss.

5. First empirical fact

- It turns out that the best prediction occurs when we use the following four characteristics:
  - the smallest daily price,
  - the largest daily price,
  - the opening price, and '
  - the closing price.
- How can we explain this empirical fact?
- In this talk, we provide a theoretical explanation for this empirical phenomenon.

6. Need to estimate the stock's beta

- Of course, skilled financial gurus do not just use computer predictions.
- They also add their knowledge and their skills.
- To best exercise this knowledge, they need to know the major characteristics of each financial instrument.
- One of the most widely used characteristic of this type is *beta* β.
- This parameter describes the linear dependence $r - r_0 \approx \beta \cdot (r_m - r_0)$, where:
  - r is return on the stock (as measured by adding the relative change in its price and the relative value of the dividends paid),
  - r_0 is the risk-free rate of return (e.g., investment in US bonds), and
  - r_m is the average market's rate of return.

7. Which values r and r_m should we use?

- If we only use the closing prices, then we have no choice.
- We use the closing price r for the individual stock and the closing price r_m for the whole market.
- However, if we take more information into account, we can use different values:
  - we can use opening prices,
  - we can use min and max prices,
  - we can use different combinations of all these prices.
- Which combination is the best? A natural idea is to select combinations that leads to the most accurate formula for r.
- For example, we can select the formula with the largest possible value of R^2.

8. Second empirical fact

- It turns out that the best is a linear combination of the four above-described stock prices: min, max, opening, and closing.
- How can we explain this empirical fact?
- In this talk, we provide a theoretical explanation for this empirical phenomenon as well.

9. Towards formulating the problem in precise terms

- We start with the prices $p_1, \ldots, p_n$ at different moments of time.
- We need to combine these prices into several characteristics.
- Different characteristics correspond to different combination rules.
- In each such rule, the combination can be done in real time.
- First, when we observe the first two prices p_1 and p_2, we combine them into a single value.
- Let us denote the result of this combination by $p_1 * p_2$.
- Then, as we observe the third value p_3, we combine the previous result with this new value.
- Thus, we get $(p_1 * p_2) * p_3$, etc.

10. Towards formulating the problem in precise terms (cont-d)

- Alternatively, if for some reason we missed the first value p_1, we could:
  - first combine p_2 and p_3 into a single value $p_2 * p_3$, and
  - then, once we learn the value p_1, combine it with our result-so-far, producing the value $p_1 * (p_2 * p_3)$.
- The combination result should reflect the stock's overall behavior.
- It should not depend on the order in which we processed the data.
- Thus, it is reasonable to expect that we have

$$(p_1 * p_2) * p_3 = p_1 * (p_2 * p_3).$$

- In mathematical terms, this means that the combination operation be *associative.*

11. The result of the combination should be within the same bounds as the combined values

- Another natural requirement is that:
  - the result $p_1 * p_2$ of combining two prices
  - should be within the same range as the original values p_1 and p_2.
- In other words, this result must be between the smallest and the largest of these two values:

$$\min(p_1, p_2) \le p_1 * p_2 \le \max(p_1, p_2).$$

12. Scale-invariance

- The result should not depend on what unit we use:
  - whether we consider prices in dollar
  - or translate them into Euros or pounds (or Thai Bahts).
- What if, instead of the original monetary unit, we use a new unit which is k times smaller.
- Then all numerical values are multiplied by k.
- So, in the new units:
  - instead of the original value p_1, we get $k \cdot p_1$,
  - instead of the original value p_2, we get $k \cdot p_2$, and
  - instead of the combined value $p_1 \cdot p_2$, we get $k \cdot (p_1 * p_2)$.
- We can combine values in the original units and then transforming to new units.

13. Scale-invariance (cont-d)

- Alternatively, we could combine the values $k \cdot p_1$ and $k \cdot p_2$ and get the result $(k \cdot p_1) * (k \cdot p_2)$.
- A natural requirement is that the combination result should not depend on what monetary units we choose, i.e.:

$$k \cdot (p_1 * p_2) = (k \cdot p_1) * (k \cdot p_2).$$

14. Shift-invariance

- As we have mentioned in our description of the beta coefficient, what is important is not so much the actual price of a stock.
- What is important is the difference $p_i - p_0$ between:
  - the stock price and
  - the value p_0 we would have gotten if we instead invested this amount in bonds.
- The bond's prices also fluctuate:
  - the change in the bond price from p_0 to a different amount $p_0 + a$ is equivalent to
  - a constant shift in all the values of the stock price, from p_i to $p_i + a$.
- Indeed, after this change, the difference remains the same:

$$(p_i + a) - (p_0 + a) = p_i - p_0.$$

15. Shift-invariance (cont-d)

- It is therefore reasonable to require that:
  - the result of the combination does not change
  - if we replace all original values p_i with shifted values $p_i + a$.
- After this replacement:
  - instead of the original value p_1, we get $p_1 + a$,
  - instead of the original value p_2, we get $p_2 + a$, and
  - instead of the combined value $p_1 \cdot p_2$, we get $(p_1 * p_2) + a$.
- We can combine the original values and then perform the shift.
- Alternative, we cane combine the shifted values $p_1 + a$ and $p_2 + a$ and get the result $(p_1 + a) * (p_2 + a)$.

16. Shift-invariance (cont-d)

- A natural requirement is that the combination result should not depend on:
  - whether we use the original values
  - or use the shifted values:

$$(p_1 * p_2) + a = (p_1 + a) * (p_2 + a).$$

- Now, we can formulate our main result.

17. Definitions

- Let $a * b$ be a binary operation that transforms pairs of real numbers into real numbers.
- We say that $*$ is *associative* if $(p_1 * p_2) * p_3 = p_1 * (p_2 * p_3)$ for all p_1, p_2, and p_3.
- We say that $*$ is *bounded* if $\min(p_1, p_2) \leq p_2 * p_2 \leq \max(p_1, p_2)$ for all p_1 and p_2.
- We say that $*$ is *scale-invariant* if $(k \cdot p_1) * (k \cdot p_2) = k \cdot (p_1 * p_2)$ for all p_1, p_2, and $k > 0$.
- We say that $*$ is *shift-invariant* if $(p_1 + a) * (p_2 + a) = p_1 * p_2 + a$ for all p_1, p_2, and a.

18. First result

- *Every associative, bounded, scale- and shift-invariant operation has one the following forms:*

$$p_1 * \ldots * p_n = \min(p_1, \ldots, p_n);$$

$$p_1 * \ldots * p_n = \max(p_1, \ldots, p_n);$$

$$p_1 * \ldots * p_n = p_1;$$

$$p_1 * \ldots * p_n = p_n.$$

- *Vice versa, each of these four operations is associative, bounded, scale- and shift-invariant.*

19. Comment

- Interestingly, a similar result can be proven for a different problem:
- How the overall emotional experience depends on the emotions experiences at different moments of time.
- In this case too, empirical data shows that the most important are the extreme and the end experiences.

20. Why Linear Combination of Four Characteristics: Explaining the Second Empirical Phenomenon

- When we combine different characteristics, it is still reasonable to require boundness and scale- and shift-invariance.
- However:
  - in contrast to the previous case when we combined similar quantities,
  - here the quantities we combine are different.
- So, in principle, we could use different combination operations for combining different characteristics.
- Thus, the associativity requirement becomes more complicated.

21. Why Linear Combination of Four Characteristics: Explaining the Second Empirical Phenomenon (cont-d)

- Also here:
  - in contrast to the previous case, while the starting price appears first,
  - all three other combined priced appear at the same time – at the end of the day.
- So there is no longer a fixed order in which we should combine these characteristics.
- We will call the corresponding version of associativity *s-associativity* (s for stock).
- Let us describe this in precise terms.

22. Definitions

- By a *combination operation*, we mean a bounded scale- and shift-invariant operation.
- We say that a function $F(c_1, c_2, c_3, c_4)$ is *s-associative* if:
  - for each permutation
$$\pi : \{1,2,3,4\} \to \{1,2,3,4\},$$
  - there exist combination operations
$$*_{\pi(1)\pi(2)}, *_{\pi(1)\pi(2)\pi(3)}, \text{ and } *_{\pi(1)\pi(2)\pi(3)\pi(4)} \text{ for which}$$
$$F(c_1,\ldots,c_4) = ((c_{\pi(1)} *_{\pi(1)\pi(2)} c_{\pi(2)}) *_{\pi(1)\pi(2)\pi(3)} c_{\pi(3)}) *_{\pi(1)\pi(2)\pi(3)\pi(4)} c_{\pi(4)}.$$

23. Definitions (cont-d)

- In other words:
  - first we combine $c_{\pi(1)}$ and $c_{\pi(2)}$ into $c_{\pi(1)} *_{\pi(1)\pi(2)} c_{\pi(2)}$;
  - then, we combine the previous result with $c_{\pi(3)}$, resulting in

$$\left(c_{\pi(1)} *_{\pi(1)\pi(2)} c_{\pi(2)}\right) *_{\pi(1)\pi(2)\pi(3)} c_{\pi(3)};$$

  - finally, we combine the previous result with $c_{\pi(4)}$.
- For example, for the trivial permutation $\pi(i) = i$, we get the following:
  - first we combine c_1 and c_2 into $c_1 *_{12} c_2$;
  - then, we combine the previous result c_3, resulting in

$$\left(c_1 *_{12} c_2\right) *_{123} c_3;$$

  - finally, we combine the previous result with c_4, resulting in

$$\left(\left(c_1 *_{12} c_2\right) *_{123} c_3\right) *_{1234} c_4.$$

24. Second result

- *Every s-associative function is a convex combination of the four characteristics* c_i.
- *Vice versa, every convex combination of the four characteristics is s-associative.*

25. Proof of the first result

- That all four operations satisfy the desired properties is easy to show.
- Let us show that, vice versa, each operation $p_1 * p_2$ that satisfies these properties has only of the four forms.
- For this, let us consider three possible relations:
  - $p_1 = p_2$,
  - $p_1 < p_2$, and
  - $p_1 > p_2$.

26. Proof of the first result: case when $p_1 = p_2$

- Let us first consider the case when $p_1 = p_2$.
- Then, boundness implies that $p_1 * p_1 = p_1$.

27. Proof of the first result: case when $p_1 < p_2$

- For $p_1 < p_2$, for $a = p_1$ and $k = p_2 - p_1$, we get $k \cdot 0 + a = p_1$ and $k \cdot 1 + a = p_2$.
- Thus, due to scale- and shift-invariance, we have

$$p_1 * p_2 = (k \cdot 0 + a) * (k \cdot 1 + a) = (k \cdot 0) * (k \cdot 1) + a = k \cdot (0 * 1) + a.$$

- Thus, $p_1 * p_2 = \alpha \cdot (p_2 - p_1) + p_1 = \alpha \cdot p_2 + (1 - \alpha) \cdot p_1$.
- Here, we denoted $\alpha \stackrel{\text{def}}{=} 0 * 1$.
- From boundess, we conclude that $0 \le \alpha = 0 * 1 \le 1$.
- Let us now use associativity.
- Due to associativity, we have

$$0 * \alpha = 0 * (0 * 1) = (0 * 0) * 1 = 0 * 1 = \alpha.$$

- Here, $0 \le \alpha$, so $0 * \alpha = \alpha \cdot \alpha + (1 - \alpha) \cdot 0 = \alpha^2$.
- From the condition $\alpha^2 = \alpha$, we conclude that either $\alpha = 0$ or $\alpha = 1$.

28. Proof of the first result: case when $p_1 < p_2$ (cont-d)

- In the first case, when $\alpha = 0$, we have $p_1 * p_2 = p_1$.
- In the second case, when $\alpha = 1$, we have $p_1 * p_2 = p_2$.

29. Proof of the first result: case when $p_1 > p_2$

- For $p_1 > p_2$, for $a = p_2$ and $k = p_1 - p_2$, we get $k \cdot 1 + a = p_1$ and $k \cdot 0 + a = p_2$.
- Thus, due to scale- and shift=invariance, we have

  $$p_1 * p_2 = (k \cdot 1 + a) * (k \cdot 0 + a) = (k \cdot 1) * (k \cdot 0) + a = k \cdot (1 * 0) + a.$$

- So, $p_1 * p_2 = \beta \cdot (p_1 - p_2) + p_2 = \beta \cdot p_1 + (1 - \beta) \cdot p_2$.
- Here we denoted $\beta \stackrel{\text{def}}{=} 1 * 0$.
- From boundness, we conclude that $0 \le \beta = 1 * 0 \le 1$.
- Let us now use associativity.
- Due to associativity, we have

  $$\beta * 0 = (1 * 0) * 0 = 1 * (0 * 0) = 1 * 0 = \beta.$$

- Here, $0 \le \beta$, so $\beta * 0 = \beta \cdot \beta + (1 - \beta) \cdot 0 = \beta^2$.
- From the condition $\beta^2 = \beta$, we conclude that either $\beta = 0$ or $\beta = 1$.

30. Proof of the first result: case when $p_1 > p_2$ (cont-d)

- In the first case, when $\beta = 0$, we have $p_1 * p_2 = p_2$.
- In the second case, when $\beta = 1$, we have $p_1 * p_2 = p_1$.

31. Proof of the first result: finalizing

- So, depending on which of the two cases holds for both possible relations $p_1 \le p_2$ and $p_2 \le p_1$, we have four cases.
- If $p_1 * p_2 = p_1$ for $p_1 < p_2$ and $p_1 * p_2 = p_2$ when $p_1 > p_2$, then, in general, $p_1 * p_2 = \min(p_1, p_2)$.
- If $p_1 * p_2 = p_2$ for $p_1 < p_2$ and $p_1 * p_2 = p_1$ when $p_1 > p_2$, then, in general, $p_1 * p_2 = \max(p_1, p_2)$.
- If $p_1 * p_2 = p_1$ for $p_1 < p_2$ and $p_1 * p_2 = p_1$ when $p_1 > p_2$, then, in general, $p_1 * p_2 = p_1$.
- If $p_1 * p_2 = p_2$ for $p_1 < p_2$ and $p_1 * p_2 = p_2$ when $p_1 > p_2$, then, in general, $p_1 * p_2 = p_2$.
- Thus, we get exactly all four combination operations.
- The result is proven.

32. Proof of the second result

- It is easy to show that:
  - every convex combination operation is
  - a combination operation in the sense of our Definition.
- Thus, every convex combination of four characteristics is s-associative.
- Vice versa, let us assume that a function $F(c_1, c_2, c_3, c_4)$ is s-associative.
- For the case when $c_1 = \min$, $c_2 = \max$, $c_3 = p_1$, and $c_4 = p_n$, we will consider two permutations: 1324 and 1423.
- In the proof of first result, we showed that each combination operation $p_1 * p_2$ is equal:
  - to one convex combination when $p_1 \leq p_2$ and
  - to another one when $p_2 \leq p_1$.
- If these convex combinations are different, then the separating line between these two convex combinations has the form $p_1 = p_2$.

33. Proof of the second result (cont-d)

- Here, $c_1 \le c_3 \le c_2$.
- Thus, $c_1 *_{13} c_3$ is a convex combination of c_1 and c_3 which is bounded from above by c_3.
- From $c_1 *_{13} c_3 \le c_3 \le c_2$, we conclude that the value $(c_1 *_{13} c_3) *_{132} c_2$ is also a convex combination of $c_1 *_{13} c_3$ and c_2.
- It is, thus, a convex combination of c_1, c_2, and c_3, i.e., has the form

$$(c_1 *_{13} c_3) *_{132} c_2 = a_1 \cdot c_1 + a_2 \cdot c_2 + a_3 \cdot c_3 \text{ for some } a_i \ge 0.$$

- For these a_i, we should have $a_1 + a_2 + a_3 = 1$.
- Thus, the function $F(c_1, \ldots, c_4)$ can be described by two convex combinations of c_i.
- If these expressions are different, then the separating line has the form

$$a_1 \cdot c_1 + a_2 \cdot c_2 + a_3 \cdot c_3 = c_4.$$

34. Proof of the second result (cont-d)

- Thus, the separating line has the form

$$a_1 \cdot c_1 + a_2 \cdot c_2 + a_3 \cdot c_3 - c_4 = 0.$$

- Similarly, from $c_1 \le c_4 \le c_2$, we conclude that the function $F(c_1, \ldots, c_4)$ can be described by two convex combinations of c_i.
- If these expressions are different, then the separating line has the form

$$a'_1 \cdot c_1 + a'_4 \cdot c_4 + a'_3 \cdot c_3 = c_2 \text{ for some } a'_i.$$

- These a'_i should add up to 1.
- Thus, the separating line has the form

$$a'_1 \cdot c_1 - c_2 + a'_3 \cdot c_3 + a'_4 \cdot c_4 = 0.$$

- The two equations for the separating line cannot describe the same set.
- Indeed, the relative signs are different.

35. Proof of the second result (cont-d)

- Thus we cannot have a separating line.
- So, the whole function $F(c_, \ldots, c_4)$ is described by a single convex combination.
- The result is proven.

36. Acknowledgments

This work was supported in part by:

- National Science Foundation grants 1623190, HRD-1834620, HRD-2034030, and EAR-2225395;
- AT&T Fellowship in Information Technology;
- program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478, and
- a grant from the Hungarian National Research, Development and Innovation Office (NRDI).