Why Quantum (Wave Probability) Models Are a Good Description of Many Non-Quantum Complex Systems, and How to Go Beyond Quantum Models

Miroslav Svítek[1], Olga Kosheleva[2], Vladik Kreinovich[2], and Thach Ngoc Nguyen[3], and Thach Ngoc Nguyen[3]

[1]Czech Technical University in Prague, svitek@fd.cvut.cz
[2]University of Texas at El Paso, El Paso, Texas 79968, USA
olgak@utep.edu, vladik@utep.edu
[3]Banking University of Ho Chi Minh City, Vietnam,
Thachnn@buh.edu.vn

1. Quantum Models Are Often a Good Description of Non-Quantum Systems

- Quantum physics has been designed to describe quantum objects.
- These are objects – mostly microscopic but sometimes macroscopic as well – that exhibit quantum behavior.
- Somewhat surprisingly, however, it turns out that quantum-type techniques can also be useful in describing non-quantum complex systems.
- For example, they describe economic systems and other systems involving human behavior.
- Why quantum techniques can help in non-quantum situations is largely a mystery.

2. Quantum Models (cont-d)

- The next natural question is related to the fact that:
  - while quantum models provide a good description of non-quantum systems,
  - this description is not perfect.
- So, a natural question: how to get a better approximation?
- In this talk, we provide answers to the above two questions.

3. Ubiquity of Multi-D Normal Distributions

- To describe the state of a complex system, we need to describe the values of some quantities $x_1, \ldots, x_n$.
- In many cases, the system consists of a large number of reasonably independent parts; in this case:
  - each of the quantities x_i describing the system
  - is approximately equal to the sum of the values that describes these parts.
- E.g., the country's trade volume is the sum of the trades performed by all its companies.
- The number of country's unemployed people is the sum of numbers from different regions, etc.

4. Multi-D Normal Distributions (cont-d)

- It is known that:
  - the distribution of the sum of a large number of independent random variables
  - is – under certain reasonable conditions – close to Gaussian (normal).
- This result is known as the *Central Limit Theorem*.
- Thus, with reasonable accuracy, we can assume that:
  - the vectors $x = (x_1, \ldots, x_n)$ formed by all the quantities that characterize the system as a whole
  - are normally distributed.

5. Let Us Simplify the Description of the Multi-D Normal Distribution

- A multi-D normal distr. is uniquely characterized:
  - by its means $\mu = (\mu_1, \ldots, \mu_n)$, $\mu_i \stackrel{\text{def}}{=} E[x_i]$, and
  - by its covariance matrix $\sigma_{ij} \stackrel{\text{def}}{=} E[(x_i - \mu_i) \cdot (x_j - \mu_j)]$.
- By observing the values x_i corresponding to different systems, we can estimate the mean values μ_i.
- Instead of the original values x_i, we can consider deviations $\delta_i \stackrel{\text{def}}{=} x_i - \mu_i$; then, $E[\delta_i] = 0$, so:
  - to fully describe the distribution of the corresponding vector $\delta = (\delta_1, \ldots, \delta_n)$,
  - it is sufficient to know the covariance matrix σ_{ij}.
- Since $E[\delta_i] = 0$, we have $\sigma_{ij} = E[\delta_i \cdot \delta_j]$.

6. For Complex Systems, with a Large Number of Parameters, a Further Simplification Is Needed

- To fully describe the distribution, we need to describe all the values of the $n \times n$ covariance matrix σ_{ij}.
- In general, an $n \times n$ matrix contains n^2 elements.
- Ssince the covariance matrix is symmetric, we only need to describe $\dfrac{n \cdot (n+1)}{2} = \dfrac{n^2}{2} + \dfrac{n}{2}$ parameters.
- Can we determine all these parameters from the observations? In general in statistics:
  - if we want to find a reasonable estimate for a parameter,
  - we need to have a certain number of observations.
- Based on N observations, we can find the value of each quantity with accuracy $\approx \dfrac{1}{\sqrt{N}}$.

7. Simplification Is Needed (cont-d)

- Thus, to be able to determine a parameter with a reasonable accuracy of 20%, we need to select N for which

$$\frac{1}{\sqrt{N}} \approx 20\% = 0.2, \text{ i.e., } N = 25.$$

- So, to find the value of one parameter, we need approximately 25 observations.
- By the same logic, for any integer k, to find the values of k parameters, we need to have $25k$ observations.
- In particular, to determine $\dfrac{n\cdot(n+1)}{2} \approx \dfrac{n^2}{2}$ parameters, we need to have $25\cdot\dfrac{n^2}{2}$ observations.
- Each fully detailed observation of a system leads to n numbers $x_1, \ldots, x_n$ and thus, to n numbers $\delta_1, \ldots, \delta_n$.

8. Simplification Is Needed (cont-d)

- So, to estimate $25 \cdot \frac{n^2}{2} = 12.5 \cdot n^2$ parameters, we need to have $12.5 \cdot n$ different systems.
- And we often do not have that many system to observe.
- For example, for a detailed analysis of a country's economy, we need to have $n \geq 30$ parameters.
- To fully describe the joint distribution of all these parameters, we need $\geq 12.5 \cdot 30 \approx 375$ countries.
- We do not have that many countries.
- This problem occurs not only in econometrics, it is even more serious, e.g., in medical bioinformatics.
- There are thousands of genes, and not enough data to be able to determine all the correlations between them.

9. Simplification Is Needed (cont-d)

- So we cannot determine the covariance matrix σ_{ij} exactly.
- Thus, we need to come up with an approximate description, with fewer parameters.

10. Need for a Geometric Description

- What does it means to have a good approximation?
- Intuitively, approximations means having a model which is, in some reasonable sense, close to the original one.
- In other words, we need model whose *distance* from the original model is small.
- So, we need to represent objects by points in a metric space.
- So, it is desirable to use an appropriate geometric representation of multi-D normal distributions.
- Such a representation is well known.

11. Geometric Description of Multi-D Normal Distribution: Reminder

- Let X be a "standard" normal distribution, with 0 mean and standard deviation 1.
- A 1D normally distributed random variable x with 0 mean and st. dev. σ can be presented as $\sigma \cdot X$.
- Similarly:
  - any normally distributed n-dimensional random vector $\delta = (\delta_1, \ldots, \delta_n)$
  - can be represented as linear combinations $\delta_i = \sum\limits_{j=1}^{n} a_{ij} \cdot X_j$ of n independent standard $X_1, \ldots, X_n$.
- The variables X_i can be found, e.g., as eigenvectors of the covariance matrix.

12. Geometric Description (cont-d)

- This way, each of the original quantities δ_i is represented by the n-dimensional vector $a_i = (a_{i1}, \ldots, a_{in})$.
- For every two linear combinations $\delta' = \sum\limits_{i=1}^{n} c'_i \cdot \delta_i$ and $\delta'' = \sum\limits_{i=1}^{n} c''_i \cdot \delta_i$ of the quantities δ_i:
  - the standard deviation $\sigma[\delta' - \delta'']$ of the difference between these linear combinations is equal to
  - the (Euclidean) distance $d(a', a'')$ between the vectors $a' = \sum\limits_{i=1} c'_i \cdot a_i$ and $a'' = \sum\limits_{i=1} c''_i \cdot a_i$, where:

$$a'_j = \sum_{i=1}^{n} c'_i \cdot a_{ij} \text{ and } a''_j = \sum_{i=1}^{n} c''_i \cdot a_{ij}.$$

13. Using Geometric Description to Find a Fewer-Parameters ($k \ll n$) Approximation

- We have n quantities $x_1, \ldots, x_n$ that describe the complex system.
- By subtracting the mean values μ_i from each of the quantities, we get shifted values $\delta_1, \ldots, \delta_n$.
- To absolutely accurately describe the joint distribution of these n quantities, we need to:
  - describe n n-dimensional vectors $a_1, \ldots, a_n$
  - corresponding to each of these quantities.
- In our approximate description, we still want to keep all n quantities.

14. Using Geometric Description (cont-d)

- However, we cannot keep them as n-dimensional vectors:
  - this would require too many parameters to determine, and
  - we do not have that many observations to be able to experimentally determine all these parameters.
- Thus, the natural thing to do is to decrease their dimension.
- In other words:
  - instead of representing each δ_i as an n-D vector $a_i = (a_{i1}, \ldots, a_{in})$ corr. to $\delta_i = \sum\limits_{j=1}^{n} a_{ij} \cdot X_j$,
  - we select some $k \ll n$ and represent each δ_i as a k-D vector $a_i = (a_{i1}, \ldots, a_{ik})$ corr. to $\delta_i = \sum\limits_{j=1}^{k} a_{ij} \cdot X_j$.

15. For $k = 2$, the Above Approximation Idea Leads to a Quantum-Type Description

- In one of the simplest cases $k = 2$, each quantity δ_i is represented by a 2-D vector $a_i = (a_{i1}, a_{i2})$.
- Similarly to the above full-dimensional case, for every two linear combinations $\delta' = \sum_{i=1}^{n} c'_i \cdot \delta_i$ and $\delta'' = \sum_{i=1}^{n} c''_i \cdot \delta_i$:
  - the standard deviation $\sigma[\delta' - \delta'']$ of the difference between these linear combinations is equal to
  - the distance $d(a', a'')$ between the corresponding 2-D vectors $a' = \sum_{i=1}^{n} c'_i \cdot a_i$ and $a'' = \sum_{i=1}^{n} c''_i \cdot a_i$, with

$$a'_j = \sum_{i=1}^{n} c'_i \cdot a_{ij} \text{ and } a''_j = \sum_{i=1}^{n} c''_i \cdot a_{ij}.$$

16. Quantum-Type Description (cont-d)

- In the 2-D case, we can alternatively represent each 2-D vector $a_i = (a_{i1}, a_{i2})$ as a complex number:

$$a_i = a_{i1} + \mathrm{i} \cdot a_{i2}.$$

- Then, the modulus $|a' - a''|$ of the difference $a' - a''$ is the distance between the original points.
- Thus, in this approximation, each quantity is represented by a complex number, and:
  - the standard deviation of the difference between different quantities is equal to
  - the modulus of the difference between the corresponding complex numbers.
- This is exactly what happens when we use quantum-type formulas.

17. Quantum-Type Description (cont-d)

- *Thus, we have indeed explained the empirical success of quantum-type formulas as a reasonable approximation.*
- Similar argument explain why, in fuzzy logic, complex-valued techniques have also been successfully used.

18. What Can We Do to Get a More Accurate Description of Complex Systems?

- As we have mentioned earlier,
  - while quantum-type descriptions are often reasonably accurate,
  - quantum formulas often do not provide the exact description of the corresponding complex systems.
- So, how can we extend and/or modify these formulas to get a more accurate description?
- Based on the above arguments, a natural way to do is:
  - to switch from complex-valued 2-dimensional ($k = 2$) approximate descriptions
  - to higher-dimensional ($k = 3$, $k = 4$, etc.) descriptions.
- Here, each quantity is represented by a k-dimensional vector.

19. A More Accurate Description (cont-d)

- The σ of each linear combination is equal to the length of the corresponding linear combination of vectors.
- For $k = 4$, we can describe this representation in terms of *quaternions* $a + b \cdot \mathrm{i} + c \cdot \mathrm{j} + d \cdot \mathrm{k}$, where:

$$\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = -1, \quad \mathrm{i} \cdot \mathrm{j} = \mathrm{k}, \quad \mathrm{j} \cdot \mathrm{k} = \mathrm{i}, \quad \mathrm{k} \cdot \mathrm{i} = \mathrm{j},$$

$$\mathrm{j} \cdot \mathrm{i} = -\mathrm{k}, \quad \mathrm{k} \cdot \mathrm{j} = -\mathrm{i}, \quad \mathrm{i} \cdot \mathrm{k} = -\mathrm{j}.$$

- For $k = 8$, we can represent it in terms of *octonions*, etc.
- Similar representations are possible for multi-D generalizations of complex-valued fuzzy logic.

20. Acknowledgments

- This work was supported:
  - by the Project AI & Reasoning CZ.02.1.01/0.0/0.0/15003/0000466,
  - by the European Regional Development Fund,
  - by the US National Science Foundation grant HRD-1242122 (Cyber-ShARE Center).
- This work was performed when M. Svítek was a Visiting Professor at the University of Texas at El Paso.
- The authors are thankful to Vladimir Mařik and Hung T. Nguyen for their support and valuable discussions.