Decision Making Under Uncertainty with Applications to Geosciences and Finance

Laxman Bokati
Computational Science Program
University of Texas at El Paso
El Paso, TX 79968, USA
lbokati@miners.utep.edu
1. Outline

- Introduction
- Decision theory: a brief reminder
- What is the optimal optimization family
- Teaching optimization
- How to speed up computations
- Applications to finance:
  - is “no trade theorem” really a paradox
  - why buying and selling prices differ
  - explaining “telescoping effect”
- Application to geosciences
- Future work plans
2. Formulation of the Problem

- In many practical situations, we need to make a decision.
- In many applications, we do not know the exact consequences of each action.
- In such situations, we need to make a decision under uncertainty.
- In many application areas, uncertainty is small – and can be made even smaller by extra measurements.
- For example, for a self-driving car, we can accurately measure all the related values and events.
- However, there are applications when it is difficult to decrease uncertainty.
- One such area is anything related to human activities.
3. Formulation of the Problem (cont-d)

- Humans make individual decisions based on their perceived value of different alternatives.
- Such behavior affects economics and finance.
- So in economics and finance, it is important to make decision under uncertainty.
- Another area where it is difficult to decrease uncertainty is geosciences.
- The only way to get a more accurate picture of what is going on beneath the earth surface is to dig a well.
- But the whole purpose of decision making is to decide whether such an expensive procedure is worth doing.
- Decision making under under uncertainty and its applications is the main topic of my research.
4. Decision Theory: A Brief Reminder

- To make a decision, we must:
  - find out the user’s preference, and
  - help the user select an alternative which is the best
    - according to these preferences.

- Traditional approach is based on an assumption that for each two alternatives $A'$ and $A''$, a user can tell:
  - whether the first alternative is better for him/her; we will denote this by $A'' < A'$;
  - or the second alternative is better; we will denote this by $A' < A''$;
  - or the two given alternatives are of equal value to the user; we will denote this by $A' = A''$. 
5. The Notion of Utility

- Under the above assumption, we can form a natural numerical scale for describing preferences.
- Let us select a very bad alternative $A_0$ and a very good alternative $A_1$.
- Then, most other alternatives are better than $A_0$ but worse than $A_1$.
- For every prob. $p \in [0, 1]$, we can form a lottery $L(p)$ in which we get $A_1$ w/prob. $p$ and $A_0$ w/prob. $1 - p$.
- When $p = 0$, this lottery simply coincides with the alternative $A_0$: $L(0) = A_0$.
- The larger the probability $p$ of the positive outcome increases, the better the result:

$$p' < p'' \text{ implies } L(p') < L(p'').$$
6. The Notion of Utility (cont-d)

- Finally, for \( p = 1 \), the lottery coincides with the alternative \( A_1 \): \( L(1) = A_1 \).
- Thus, we have a continuous scale of alternatives \( L(p) \) that monotonically goes from \( L(0) = A_0 \) to \( L(1) = A_1 \).
- Due to monotonicity, when \( p \) increases, we first have \( L(p) < A \), then we have \( L(p) > A \).
- The threshold value is called the utility of the alternative \( A \):
  \[
  u(A) \overset{\text{def}}{=} \sup\{p : L(p) < A\} = \inf\{p : L(p) > A\}.
  \]
- Then, for every \( \varepsilon > 0 \), we have
  \[
  L(u(A) - \varepsilon) < A < L(u(A) + \varepsilon).
  \]
- We will describe such (almost) equivalence by \( \equiv \), i.e., we will write that \( A \equiv L(u(A)) \).
7. A Rational Agent Should Maximize Utility

- Suppose that we have found the utilities \( u(A') \), \( u(A'') \), \( \ldots \), of the alternatives \( A', A'', \ldots \)
- Which of these alternatives should we choose?
- By definition of utility, we have:
  - \( A \equiv L(u(A)) \) for every alternative \( A \), and
  - \( L(p') < L(p'') \) if and only if \( p' < p'' \).
- We can thus conclude that \( A' \) is preferable to \( A'' \) if and only if \( u(A') > u(A'') \).
- In other words, we should always select an alternative with the largest possible value of utility.
8. How to Estimate Utility of an Action

- For each action, we usually know possible outcomes $S_1, \ldots, S_n$.
- We can often estimate the prob. $p_1, \ldots, p_n$ of these outcomes.
- By definition of utility, each situation $S_i$ is equiv. to a lottery $L(u(S_i))$ in which we get:
  - $A_1$ with probability $u(S_i)$ and
  - $A_0$ with the remaining probability $1 - u(S_i)$.
- Thus, the action is equivalent to a complex lottery in which:
  - first, we select one of the situations $S_i$ with probability $p_i$: $P(S_i) = p_i$;
  - then, depending on $S_i$, we get $A_1$ with probability $P(A_1 | S_i) = u(S_i)$ and $A_0$ w/probability $1 - u(S_i)$. 
9. How to Estimate Utility of an Action (cont-d)

• Reminder:
  • first, we select one of the situations $S_i$ with probability $p_i$: $P(S_i) = p_i$;
  • then, depending on $S_i$, we get $A_1$ with probability $P(A_1 | S_i) = u(S_i)$ and $A_0$ w/probability $1 - u(S_i)$.

• The prob. of getting $A_1$ in this complex lottery is:
  \[
P(A_1) = \sum_{i=1}^{n} P(A_1 | S_i) \cdot P(S_i) = \sum_{i=1}^{n} u(S_i) \cdot p_i.
  \]

• In the complex lottery, we get:
  • $A_1$ with prob. $u = \sum_{i=1}^{n} p_i \cdot u(S_i)$, and
  • $A_0$ w/prob. $1 - u$.

• So, we should select the action with the largest value of expected utility $u = \sum p_i \cdot u(S_i)$. 
10. To Practical Applications of Decision Theory

- The numerical value of utility depends on the selection of the alternatives $A_0$ and $A_1$.
- If we select a different pair $(A'_0, A'_1)$, then utility changes into $u'(A) = a \cdot u(A) + b$ for some $a > 0$ and $b$.
- The dependence of utility of money is non-linear.
- Utility $u$ is proportional to the square root of the amount $m$ of money $u = c \cdot \sqrt{m}$.
- If we have an amount $m$ of money now, then we can place it in a bank and add an interest.
- So, we get the new amount $m' \overset{\text{def}}{=} (1 + i) \cdot m$ in a year.
- Thus, the amount $m'$ in a year is equivalent to the value $m = q \cdot m'$ now, where $q \overset{\text{def}}{=} 1/(1 + i)$.
- This is called discounting.
11. Decision Making under Interval Uncertainty

- In real life, we rarely know the exact consequences of each action.

- So, for an alternative $A$, we often only know the bounds on $u(A)$: $\underline{u}(A) \leq u(A) \leq \overline{u}(A)$.

- For such an interval case, we need to be able to compare the interval-valued alternative with lotteries $L(p)$.

- As a result of such comparison, we will come up with a utility of this interval.

- So, we need to assign, to each interval $[\underline{u}, \overline{u}]$, a utility value $u(\underline{u}, \overline{u}) \in [\underline{u}, \overline{u}]$.

- Reminder: utility is determined modulo a linear transformation $u' = a \cdot u + b$. 
12. Case of Interval Uncertainty (cont-d)

- *Reasonable to require:* the equivalent utility does not change with re-scaling: for \( a > 0 \) and \( b \),

\[ u(a \cdot u^- + b, a \cdot u^+ + b) = a \cdot u(u^-, u^+) + b. \]

- For \( u^- = 0, u^+ = 1, a = \bar{u} - \underline{u}, \) and \( b = \underline{u} \), we get

\[ u(\underline{u}, \bar{u}) = \alpha_H \cdot (\bar{u} - \underline{u}) + \underline{u} = \alpha_H \cdot \bar{u} + (1 - \alpha_H) \cdot \underline{u}. \]

- This formula was first proposed by a future Nobelist Leo Hurwicz.

- It is known as the Hurwicz optimism-pessimism criterion.
13. Which Is the Optimal Approximation Family?

- We need to approximate the actual dependence as \( f(x) = \sum_{i=1}^{n} C_i \cdot f_i(x) \), where \( f_i(x) \) are given functions.

- A reasonable requirement is related to the fact that the numerical value of \( x \) depends:
  - on the choice of a measuring unit (years or months),
  - and on the choice of a starting point.

- If we change a measuring unit by a new one which is \( a \) times smaller, then \( x \rightarrow a \cdot x \).

- If we replace the original starting point with the new one, \( b \) units in the past \( x \rightarrow x + b \).

- The general formulas for extrapolation should not depend on selecting a unit or a starting point.
14. Choosing $f_i(x)$ (cont-d)

• It is therefore reasonable to assume that the approximating family \( \left\{ \sum_{i=1}^{n} C_i \cdot f_i(x) \right\}_{C_1,\ldots,C_n} \) will not change:

\[
\left\{ \sum_{i=1}^{n} C_i \cdot f_i(a \cdot x) \right\}_{C_1,\ldots,C_n} = \left\{ \sum_{i=1}^{n} C_i \cdot f_i(x + b) \right\}_{C_1,\ldots,C_n} = \left\{ \sum_{i=1}^{n} C_i \cdot f_i(x) \right\}_{C_1,\ldots,C_n}.
\]

• It turns out that under these conditions, all the basic functions are polynomials.

• So, all their linear combinations are polynomials.

• Thus, it is reasonable to approximate functions by polynomials.
15. Proof: Main Ideas

- Shift-invariance implies that for each $i$, we have

$$f_i(x + x_0) = \sum_{j=1}^{n} C_{ij}(x_0) \cdot f_j(x).$$

- Differentiating w.r.t. $x_0$ and taking $x_0 = 0$, we get

$$f'_i(x) = \sum_{j=1}^{n} c_{ij} \cdot f_j(x),$$

where $c_{ij} \overset{\text{def}}{=} C'_{ij}(x)$. 

- Solutions are such system are known.

- They are linear combinations of functions $x^k \cdot \exp(\lambda_k \cdot x)$, where $k$ is a natural number and $\lambda_k$ is complex.

- Scale-invariance means $f_i(\lambda \cdot x) = \sum_{j=1}^{n} C_{ij}(\lambda) \cdot f_j(x)$.

- Differentiating w.r.t. $\lambda$ and taking $\lambda = 1$, we get

$$x \cdot f'_i(x) = \sum_{j=1}^{n} c_{ij} \cdot f_j(x).$$
16. Proof: Main Ideas (cont-d)

- For \( X = \ln(x) \) and \( F_i(X) = f_i(\exp(X)) \), we get

\[
F_i'(x) = \sum_{j=1}^{n} c_{ij} \cdot F_j(X).
\]

- Thus, \( F_i(X) \) is a linear combination of \( X^k \cdot \exp(\lambda_k \cdot X) \).

- Hence, \( f_i(x) = F_i(\ln(x)) \) is a linear combination of terms \( (\ln(x))^k \cdot x^\lambda \).

- One can easily see that the only functions common to both families are polynomials.
17. Teaching Optimization: How to Generate “Nice” Cubic Polynomials

- We have shown that it is reasonable to approximate functions by polynomials.
- It is therefore important to teach decision makers how to optimize such functions.
- In general, people feel more comfortable with rational numbers than with irrational ones.
- Thus, it is desirable to have examples of simple problems for which zeros and extrema points are rational.
- For quadratic functions, no calculus is needed.
- Thus, the simplest case is cubic functions.
18. Good News

• Good news is that when we know that the roots are rational, it is (relatively) easy to find these roots.

• Indeed, for each rational root \( x = \frac{p}{q} \) of a polynomial \( a_n \cdot x^n + a_{n-1} \cdot x^{n-1} + \ldots + a_0 \) with integer coefficients:
  - the numerator \( p \) is a factor of \( a_0 \), and
  - the denominator \( q \) is a factor of \( a_n \).

• Extreme points are roots of quadratic equations – also easy to find.

• So, it is sufficient to find “nice” polynomials, we can then compute roots and extreme points.

• How can we find polynomials for which both zeros and extreme points are rational?
19. What Is Known and What We Do

- An algorithm for generating such polynomials was recently proposed.
- This algorithm, however, is not the most efficient one.
- For each tuple of the corresponding parameter values, it uses exhaustive trial-and-error search.
- In this presentation, we produce an efficient algorithm for producing nice polynomials.
- Namely, we propose simple computational formulas:
  - for each tuple of the corresponding parameters, these formulas produce a “nice” cubic polynomial;
  - every “nice” cubic polynomial can be thus generated.
- For each tuple, our algorithm requires the same constant number of elementary steps.
20. Analysis of the Problem

- A general cubic polynomial with rational coefficients has the form $a \cdot X^3 + b \cdot X^2 + c \cdot X + d$.

- Roots and extreme points of $f(x)$ do not change if we simply divide all its values by the same constant $a$.

- Thus, it is sufficient to consider polynomials with only three parameters: $X^3 + p \cdot X^2 + q \cdot X + r$, where

$$ p \overset{\text{def}}{=} \frac{b}{a}, \quad q \overset{\text{def}}{=} \frac{c}{a}, \quad r \overset{\text{def}}{=} \frac{d}{a}. $$

- We can further simplify the problem if we replace $X$ with $x \overset{\text{def}}{=} X + \frac{p}{3}$, then we get $x^3 + \alpha \cdot x + \beta$, where

$$ \alpha = q - \frac{p^2}{3} \quad \text{and} \quad \beta = r - \frac{p \cdot q}{3} + \frac{2p^3}{27}. $$
21. Analysis of the Problem (cont-d)

- Let $r_1$, $r_2$, and $r_3$ denote rational roots of $x^3 + \alpha \cdot x + \beta$, then, we have

\[ x^3 + \alpha \cdot x + \beta = (x - r_1) \cdot (x - r_2) \cdot (x - r_3). \]

- So, $r_1 + r_2 + r_3 = 0$, $\alpha = r_1 \cdot r_2 + r_2 \cdot r_3 + r_1 \cdot r_3$, and $\beta = -r_1 \cdot r_2 \cdot r_3$.

- Substituting $r_3 = -(r_1 + r_2)$ into these formulas, we get

\[ \alpha = -(r_1^2 + r_1 \cdot r_2 + r_2^2) \text{ and } \beta = r_1 \cdot r_2 \cdot (r_1 + r_2). \]
22. Using the Fact That the Extreme Points $x_0$ Should Also Be Rational

- Differentiating and equating the derivative to 0, we get
  
  \[3x_0^2 - (r_1^2 + r_1 \cdot r_2 + r_2^2) = 0.\]

- This is equivalent to $3x_0^2 - 3y^2 - z^2 = 0$, where
  
  \[y \overset{\text{def}}{=} \frac{r_1 + r_2}{2} \quad \text{and} \quad z \overset{\text{def}}{=} \frac{r_1 - r_2}{2}.\]

- If we divide both sides of this equation by $y^2$, we get
  
  \[3X_0^2 - 3 - Z^2 = 0, \quad \text{where} \quad X_0 \overset{\text{def}}{=} \frac{x_0}{y} \quad \text{and} \quad Z \overset{\text{def}}{=} \frac{z}{y}.\]

- One of the solution of above equation is easy to find: namely, when $X_0 = -1$, we get $Z^2 = 0$ and $Z = 0$.

- This means that for every $y$, $x_0 = -y$, $y$ and $z = 0$ solve the above equation.
23. Using the Fact That the Extreme Points $x_0$ Should Also Be Rational (cont-d)

- We can now reconstruct $r_1$ and $r_2$ from $y$ and $z$ as $r_1 = y + z$ and $r_2 = y - z$.
- In our case, $r_1 = r_2 = y$, so $\alpha = -3y^2$ and $\beta = 2y^3$.
- We can then:
  - shift by a rational number $s$ ($x \rightarrow X + s$), and
  - multiply all the coefficients by an arbitrary rational number $a$.
- Then, we get
  
  \[ b = 3a \cdot s, \quad c = a \cdot (3s^2 - 3y^2), \quad d = a \cdot (s^3 + 2y^3). \]
24. Using the General Algorithm for Finding All Rational Solutions to a Quadratic Equation

• We have already found a solution of the equation $3X_0^2 - 3 - Z^2 = 0$, corresponding to $X_0 = -1$: then $Z = 0$.

• Every other solution $(X_0, Z)$ can be connected to this simple solution $(-1, 0)$ by a straight line.

• A general equation of a straight line passing through the point $(-1, 0)$ is $Z = t \cdot (X_0 + 1)$.

• When $X_0$ and $Z$ are rational, $t = \frac{Z}{X_0 + 1}$ is rational.

• Substituting this expression for $Z$ into the equation, we get $3X_0^2 - 3 - t^2 \cdot (X_0 + 1)^2 = 0$.

• Since $X_0 \neq -1$, we can divide both sides by $X_0 + 1$. then $3 \cdot (X_0 - 1) - t^2 \cdot (X_0 + 1) = 0$, hence

$$X_0 = \frac{3 + t^2}{3 - t^2} \quad \text{and} \quad Z = \frac{6t}{3 - t^2}.$$
Towards a General Description of All “Nice” Polynomials

• For every rational $y$, we can now take $x_0 = y \cdot X_0$, $y$, and $z = y \cdot Z = \frac{6yt}{3 - t^2}$.

• Based on $y$ and $z$, we can compute $r_1 = y + z$ and $r_2 = y - z$.

• Then, we can compute $\alpha$ and $\beta$:
  
  $$\alpha = -3y^2 - z^2 \text{ and } \beta = 2y \cdot (y^2 - z^2).$$

• Now, we can apply shift by $s$ and multiplication by $a$.

• Thus, we arrive at the following algorithm for computing all possible “nice” cubic polynomials.
26. Resulting Algorithm for Computing All “Nice” Cubic Polynomials

• We use four arbitrary rational numbers \( t, y, s, \) and \( a; \) based on these numbers, we first compute \( z = \frac{6yt}{3 - t^2}. \)

• Then, we compute the coefficients \( b, c, \) and \( d \) of the resulting “nice” polynomial (\( a \) we already know):

\[
b = 3a \cdot s; \quad c = a \cdot (3s^2 - 3y^2 - z^2); \\
d = a \cdot (s^3 + 2y \cdot (y^2 - z^2)).
\]

• These expressions cover almost all “nice” polynomials, with the exception of the following family:

\[
b = 3a \cdot s, \quad c = a \cdot (3s^2 - 3y^2), \quad d = a \cdot (s^3 + 2y^3).
\]
27. How to Speed Up Computations: Functions in Quantum and Reversible Computing

- Decision making means optimization.
- How can we solve the corresponding computational problems?
- For complex problems, the existing algorithms take too long.
- It is therefore necessary to explore the possibility of faster computations.
- One such possibility is the use of quantum computing.
- Existing algorithms do not adequately represent functions.
- We thus need to come up with more adequate representation of generic functions in quantum computing.
28. Why Quantum Computing

- According to modern physics, all processes cannot move faster than the speed of light.
- For a laptop of size $\approx 30$ cm, it takes at least $1$ nanosecond ($10^{-9}$ sec) for a signal to get across.
- During this time, even the cheapest laptops perform several operations.
- Thus, to speed up computations, we need to further decrease the size of the computer.
- So, we must decrease the size of memory cells.
- Their size is already compatible with a molecule.
- If we decrease the computer cells even more, they will consist of a few dozen molecules.
- Thus, we will need to take into account the physics that describe such micro-objects – quantum physics.
29. Need for Reversible Computing

- For macro-objects, we can observe irreversible processes.
- If we drop a china cup on a hard floor, it will break into pieces.
- No physical process can combine these pieces back into the original whole cup.
- However, on the micro-level, all the equations are reversible.
- This is true for Newton’s equations, this is true for quantum Schroedinger’s equations.
- Reversible computing is also needed for a different reasons.
- Theoretically, we could place more memory cells if we stack them in 3-D.
- However, that will melt the computer.
30. Need for Reversible Computing

• The heat is caused by the Second Law of Thermodynamics:
  – every time we have an irreversible process,
  – heat is radiated, in the amount $T \cdot S$, where $S$ is the entropy,
  – in this case, $S$ is the number of bits in information loss.

• Basic logic operations (that underlie all computations) are irreversible.

• For example, when $a \& b$ is false, we cannot uniquely determine $a$ and $b$.

• So, to decrease the amount of heat, we should use only reversible operations.
31. How to Make Operations Reversible?

- For a bit-valued function $y = f(x_1, \ldots, x_n)$ quantum version is:
  $$T_f : (x_1, \ldots, x_n, x_0) \rightarrow (x_1, \ldots, x_n, x_0 \oplus f(x_1, \ldots, x_n)).$$
- It is indeed reversible: $T_f(T_f(x)) = x$.
- The problem is we usually compose an algorithm as a composition (successive use) of subroutines $f \circ \ldots \circ g$.
- However, $T_f \circ T_g \neq T_{f \circ g}$.
- We want to have $S_f$ for which $S_{f \circ g} = S_f \circ S_g$.
- Our recommendation is $S_f(x_1, \ldots, x_n, u) =
  \left( f_1(x_1, \ldots, x_n), \ldots, f_n(x_1, \ldots, x_n), u \middle/ \det \left\| \frac{\partial f_i}{\partial x_j} \right\| \right)$.
32. **First Application to Finance: Is “No Trade Theorem” Really a Paradox**

- One of the challenges in foundations of finance is the so-called “no trade theorem” paradox:
  - if a trader wants to sell a stock, he/she believes that this stock will go down;
  - however, another trader is willing to buy it;
  - this means that this other expert believes that the stock will go up.

- The fact that equally good experts have different beliefs should dissuade the first expert from selling.

- Thus, trades should be very rare.

- However, in reality, trades are ubiquitous; how can we explain this?
33. Our Explanation

- Let $s$ be the current cost of the stock. Let $m$ be the mean and $\sigma$ st. dev. of the (discounted) future gain $g$.
- Let $M$ be the person’s initial amount of money.
- Buying a stock is beneficial if it increases the expected utility, i.e., if $E[\sqrt{M - s + g}] > \sqrt{M}$.
- For small $s$, this is equivalent to $M > M_0 \overset{\text{def}}{=} \frac{(m - s)^2 + \sigma^2}{2(m - s)}$.
- So, folks with $M > M_0$ benefit from buying it.
- People with $M < M_0$ benefit from selling it.
- This explains the ubiquity of trading.
- The larger the risk $\sigma$, the larger the threshold $M_0$.
- This explains why depressed people (with lower equivalent value of $M = u^2$) are more risk-averse.
34. 2nd Application to Finance: Why Prices for Buying and Selling Objects Are Different

- Intuitively, we should decide, for ourselves, how much each object is worth to us.

- This worth amount should be the largest amount that we should be willing to pay if we are buying this object.

- This same amount should be the smallest amount for which we should agree to sell this object.

- However, in practice, the buying and selling prices are different.

- The main reason is that people are not clear on the value of each object.

- At best, they have a range $[u, \bar{u}]$ of possible values of this object’s worth.
35. Why Buying and Selling Prices Differ

• According to Hurwicz formula, when we buy, we gain the value $u_b = \alpha_H \cdot \bar{u} + (1 - \alpha_H) \cdot \underline{u}$.

• On the other hand, if we already own this object and we sell it, then our loss is between $-\bar{u}$ and $-\underline{u}$.

• The Hurwicz criterion estimates the resulting value as $-u_s$, where $u_s = \alpha_H \cdot \bar{u} + (1 - \alpha_H) \cdot \underline{u}$.

• In the general case, the values $u_b$ and $u_s$ are indeed different.
36. Explaining “Telescoping Effect” – That Time Perception Is Biased

- People usually underestimate time passed since distant events, and overestimate for recent events.
- Time $t$ is related to utility via discounting: $u = u_0 \cdot q^t$.
- This utility value is always in $[0, u_0]$.
- We only know utility $u$ with some accuracy $\varepsilon$.
- Instead of the original value $u = u_0 \cdot q^t$, we only know that $u \in [u_0 \cdot q^t - \varepsilon, u_0 \cdot q^t + \varepsilon]$.
- For small $t$, $u_0 \cdot q^t \approx u_0$, so $u_0 \cdot q^t + \varepsilon > u_0$.
- Thus, we have the interval $[u_0 \cdot q^t - \varepsilon, u_0]$, and Hurwicz method leads to the value
  
  $$u(t) = \alpha_H \cdot u_0 + (1 - \alpha_H) \cdot u_0 \cdot (q^t - \varepsilon).$$

- For $t \to 0$, $u_0 \cdot q^t \to u_0$ while $u(t) \to q_0 - (1 - \alpha_H) \cdot \varepsilon < u_0$. 
37. Explaining “Telescoping Effect” (cont-d)

- Thus, for small $t$, we have $u(t) < u_0 \cdot q^t$.
- The perceived time $\tilde{t}$ comes from $u(t) = u_0 \cdot \tilde{q}^t$, so $\tilde{t} > t$.
- For large $t$, we have $u_0 \cdot q^t - \varepsilon < 0$, so $u \in [0, u_0 \cdot q^t + \varepsilon]$.
- Hurwicz methods leads to the value $u(t) = \alpha_H \cdot (u_0 \cdot q^t + \varepsilon)$.
- For $t \rightarrow \infty$, $u_0 \cdot q^t \rightarrow 0$ while $u(t) \rightarrow \alpha_H \cdot \varepsilon > 0$.
- Thus, for large $t$, we have $u(t) > u_0 \cdot q^t$.
- The perceived time $\tilde{t}$ comes from $u(t) = u_0 \cdot \tilde{q}^t$, so $\tilde{t} < t$.
- This explains the telescoping effect.
38. **Application to Geosciences: Bhutan Landscape Anomaly Explained**

- In the Bhutan area of the Himalayas region, there seems to be a landscape anomaly.

- We can plot elevation profile centered at the lowest point (usually the river).

- In most of the world, the elevation profile is:
  - first convex (in the river valley), and
  - then becomes concave – which corresponds to the mountain peaks.

- In contrast, in Bhutan, the profile turns concave very fast, way before we reach the mountain peaks area.

- As of now, there are no good well-accepted explanations for this phenomenon – which makes it an anomaly.
39. Bhutan Landscape Anomaly (cont-d)

- To be more precise, we know that the geophysics of the Bhutan area is somewhat different:
  - in Nepal, the advancing tectonic plate in orthogonal to the border of the mountain range;
  - similarly, in the rest of the world;
  - in Bhutan, the plate pushes the range at an angle.
- How this explain the Bhutan anomaly?
40. Our Explanation

- Earlier we explained that polynomials are a reasonable approximation family.
- For constant and linear functions, we do not have any landscape.
- So, the simplest are quadratic functions.
- It makes sense to call \( y = f(x) \) and \( y = g(x) \) equivalent if they differ only by re-scaling and shift of \( x \) and \( y \):

\[
g(x) = \lambda_y \cdot f(\lambda_x \cdot x + x_0) + y_0.
\]

- One can show that every non-linear quadratic function is equivalent either to \( x^2 \) or to \( -x^2 \).
- So, in this approximation, we have, in effect, two shapes: \( x^2 \) (convex) and the \( -x^2 \) (concave).
41. Our Explanation (cont-d)

- This result explains why the usual visual classification into convex and concave shapes.
- To get a more accurate description, let us also consider cubic terms:
  \[ f(x) = a_0 + a_1 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3. \]
- As a starting point \( x = 0 \), we can take the lowest (or the highest) point.
- In both cases, \( f'(0) = 0 \), so \( a_1 = 0 \) and
  \[ f(x) = a_0 + a_2 \cdot x^2 + a_3 \cdot x^3. \]
- In the case of Nepal, the forces compressing the upper plate are orthogonal to the line of contact.
- This means that in this case, the forces do not change if we swap left and right: \( x \leftrightarrow -x \).
42. Our Explanation (cont-d)

• The whole mountain range was created by this force.

• So, it is reasonable to conclude that the elevation profile is also invariant w.r.t. \( x \leftrightarrow -x \): \( f(x) = f(-x) \).

• This leads to \( a_3 = 0 \).

• Thus, in this case, the elevation profile is quadratic even in this next approximation.

• It is, therefore, either convex or concave.

• In the case of Bhutan, the force is applied at an angle.

• Here, there is no symmetry with respect to \( x \rightarrow -x \), so, in general, we have \( a_3 \neq 0 \).
43. Our Explanation (cont-d)

• In general, the second derivative $f''(x)$ describes whether a function is:
  
  – locally convex (when $f''(x) > 0$) or
  – locally concave (when $f''(x) < 0$).

• In our case, $f''(x) = 6a_3 \cdot x + 2a_2$, with $a_3 \neq 0$.

• A non-constant linear function always changes signs.

• This explains why in the case of Bhutan, convexity follows by concavity.
44. Future Work Plans

• This thesis is devoted to:
  – a practically important problem of decision making under uncertainty and
  – its applications to finances and geosciences.

• In this thesis, we provide our preliminary results and first applications.

• We plan to continue this work, both:
  – by expanding our theoretical analysis and
  – by coming up with more examples of practical applications.
45. Future Plans: Theory

- In terms of theoretical analysis, what we have done so far is based on deterministic decision making.
- In practice, our decisions are often probabilistic.
- In the same situation, we may select different alternatives, with different probabilities.
- This situation has been analyzed in decision theory by a Nobelist D. McFadden.
- However, his analysis assumes that we know the exact gains related to different alternatives.
- In practice, we usually know the expected gains only with some uncertainty.
- So, our main theoretical research would be to extend McFadden’s analysis to the case of uncertainty.
46. Future Plans: Theory (cont-d)

- This planned research is closely related to the need to speed up computations.
- At present, in many problems, deep machine learning is effectively used to extract dependencies from data.
- Interestingly, one of the stages of deep learning – softmax – uses the same formulas as McFadden.
- So, it assumes that we know the exact expected gains of different possible decisions.
- We hope that accounting for uncertainty will speed up deep learning.
47. Future Plans: Explanations

- First, there are still seemingly counterintuitive aspects of human behavior that need explaining; e.g.:
  - an often cited phrase that giving is better than receiving
  - seems to be inconsistent with the usual utilitarian models of this behavior.
- Second, the Hurwicz analysis does not explain why some people are more optimistic.
- It is therefore desirable to try to understand this.
- For this purpose, we will analyze which type of behavior works best in different situations.
48. Future Plans: Explanations (cont-d)

- Finally, it is desirable to look:
  - not just at the results of human decision making,
  - but also at procedures that humans use to reach their results.

- For example, as part of these procedures, humans perform some non-traditional approximate computations.

- We plan to analyze how these unusual procedures can be explained by decision making under uncertainty.
49. Future Plans: Applications

• First, we plan to analyze how all this can be applied to grading student papers.

• As of now, deciding which problems and which tests are worth how many points is more of an art.

• This problem is difficult to solve in precise terms because we need to make decisions under uncertainty.

• We plan to see what recommendations can be extracted from decision making under uncertainty.

• Second, we plan to see if we can make robots behavior more human-like and thus, more acceptable to users.

• Finally, we will continue to look for possible challenges and applications in geosciences.
50. Acknowledgments

I want to thank:

- my advisor Dr. Vladik Kreinovich;
- my committee members Drs. Martine Ceberio and Aaron Velasco;
- my friends Pawan Koirala, Kamal Nyaupane, Neelam Dumre, Bibek Aryal, and Subharaj Ranabhat; and
- my father Damber Dutta Bokati, my mother Ishwori Devi Bokati, and my brother Prakash Bokati.

Without your inspiration and support, I would not have been able to pursue my academic goals.