28th European Conference on Operational Research

Algorithmic Future of Operations Research

Yurii Nesterov (CORE/INMA, UCL)

July 3, 2016

Predicting the future of Optimization

Predicting the future of Optimization

- Boris Polyak (1983):

Predicting the future of Optimization

- Boris Polyak (1983): "It seems that the period of highest activity in Optimization is over."

Predicting the future of Optimization

- Boris Polyak (1983): "It seems that the period of highest activity in Optimization is over."
- 1984: Interior-Point Methods

Predicting the future of Optimization

- Boris Polyak (1983): "It seems that the period of highest activity in Optimization is over."
- 1984: Interior-Point Methods
- 2003: Renaissance of Fast Gradient Methods

Predicting the future of Optimization

- Boris Polyak (1983): "It seems that the period of highest activity in Optimization is over."
- 1984: Interior-Point Methods
- 2003: Renaissance of Fast Gradient Methods
- Starting from 2010: Big Data revolution.

Predicting the future of Optimization

- Boris Polyak (1983): "It seems that the period of highest activity in Optimization is over."
- 1984: Interior-Point Methods
- 2003: Renaissance of Fast Gradient Methods
- Starting from 2010: Big Data revolution.

Observations:

Predicting the future of Optimization

- Boris Polyak (1983): "It seems that the period of highest activity in Optimization is over."
- 1984: Interior-Point Methods
- 2003: Renaissance of Fast Gradient Methods
- Starting from 2010: Big Data revolution.

Observations:

- Traditional methods are not efficient in Huge Dimension

Predicting the future of Optimization

- Boris Polyak (1983): "It seems that the period of highest activity in Optimization is over."
- 1984: Interior-Point Methods
- 2003: Renaissance of Fast Gradient Methods
- Starting from 2010: Big Data revolution.

Observations:

- Traditional methods are not efficient in Huge Dimension
- We need to learn the way Optimization is incorporated in Nature.

Hints from the Wizards

Hints from the Wizards

Leonard Euler (1744):

Hints from the Wizards

Leonard Euler (1744):

"Nothing in the world takes place without optimization, and there is no doubt that all aspects of the world that have a rational basis can be explained by optimization methods".

Hints from the Wizards

Leonard Euler (1744):

"Nothing in the world takes place without optimization, and there is no doubt that all aspects of the world that have a rational basis can be explained by optimization methods".

Adam Smith (1776):

Hints from the Wizards

Leonard Euler (1744):

"Nothing in the world takes place without optimization, and there is no doubt that all aspects of the world that have a rational basis can be explained by optimization methods".

Adam Smith (1776): *Invisible hand of the market*.

Hints from the Wizards

Leonard Euler (1744):

"Nothing in the world takes place without optimization, and there is no doubt that all aspects of the world that have a rational basis can be explained by optimization methods".

Adam Smith (1776): *Invisible hand of the market*.

Plutarch/.../Erasmus(1500):

Hints from the Wizards

Leonard Euler (1744):

"Nothing in the world takes place without optimization, and there is no doubt that all aspects of the world that have a rational basis can be explained by optimization methods".

Adam Smith (1776): *Invisible hand of the market*.

Plutarch/.../Erasmus(1500): *"God's mill grinds slow but sure"*.

Hints from the Wizards

Leonard Euler (1744):

"Nothing in the world takes place without optimization, and there is no doubt that all aspects of the world that have a rational basis can be explained by optimization methods".

Adam Smith (1776): *Invisible hand of the market*.

Plutarch/.../Erasmus(1500): *"God's mill grinds slow but sure"*.

Our guesses:

Hints from the Wizards

Leonard Euler (1744):

"Nothing in the world takes place without optimization, and there is no doubt that all aspects of the world that have a rational basis can be explained by optimization methods".

Adam Smith (1776): *Invisible hand of the market*.

Plutarch/.../Erasmus(1500): *"God's mill grinds slow but sure"*.

Our guesses:

1. Optimization algorithms are deeply involved in Nature/Social Life.

Hints from the Wizards

Leonard Euler (1744):

"Nothing in the world takes place without optimization, and there is no doubt that all aspects of the world that have a rational basis can be explained by optimization methods".

Adam Smith (1776): *Invisible hand of the market*.

Plutarch/.../Erasmus(1500): *"God's mill grinds slow but sure"*.

Our guesses:

1. Optimization algorithms are deeply involved in Nature/Social Life.

2. Very often, they are implemented as unintentional/subconscious actions.

Hints from the Wizards

Leonard Euler (1744):

"Nothing in the world takes place without optimization, and there is no doubt that all aspects of the world that have a rational basis can be explained by optimization methods".

Adam Smith (1776): *Invisible hand of the market*.

Plutarch/.../Erasmus(1500): *"God's mill grinds slow but sure"*.

Our guesses:

1. Optimization algorithms are deeply involved in Nature/Social Life.

2. Very often, they are implemented as unintentional/subconscious actions.

3. Their rate of convergence is slow.

Hints from the Wizards

Leonard Euler (1744):

"Nothing in the world takes place without optimization, and there is no doubt that all aspects of the world that have a rational basis can be explained by optimization methods".

Adam Smith (1776): *Invisible hand of the market*.

Plutarch/.../Erasmus(1500): *"God's mill grinds slow but sure"*.

Our guesses:

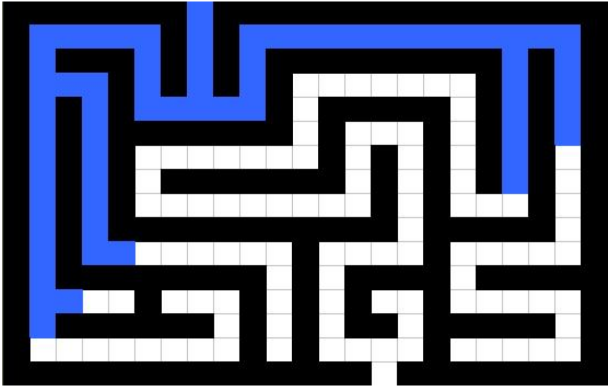
1. Optimization algorithms are deeply involved in Nature/Social Life.

2. Very often, they are implemented as unintentional/subconscious actions.

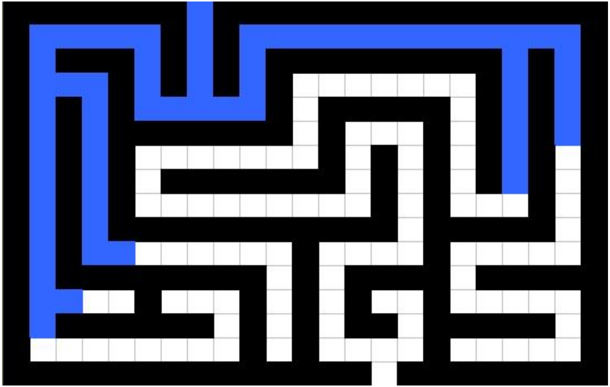
3. Their rate of convergence is slow. However, they have reasonable worst-case performance guarantees.

Example I: Flood Dynamics

Example I: Flood Dynamics

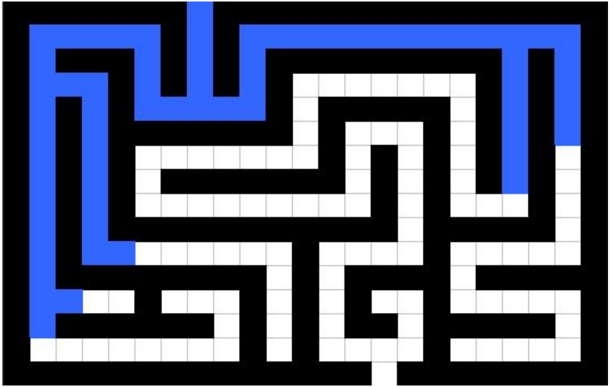


Example I: Flood Dynamics



Theorem: Every point is reached in minimal time.

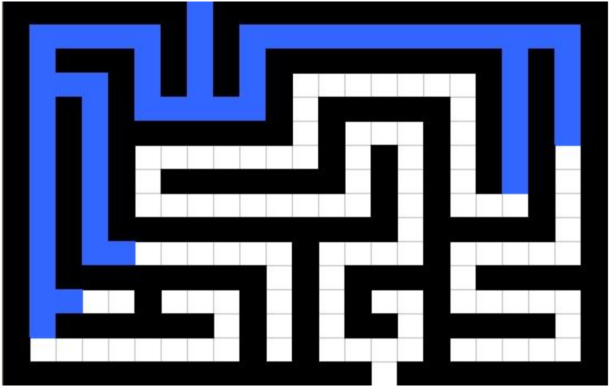
Example I: Flood Dynamics



Theorem: Every point is reached in minimal time.

Other examples:

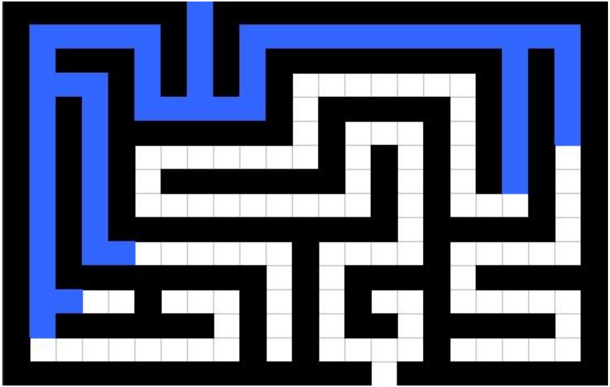
Example I: Flood Dynamics



Theorem: Every point is reached in minimal time.

Other examples: light,

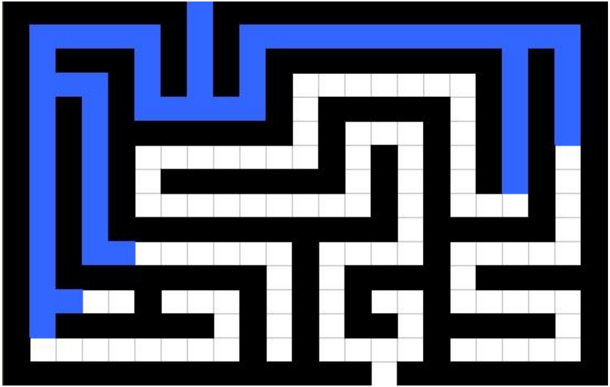
Example I: Flood Dynamics



Theorem: Every point is reached in <u>minimal time</u>.

Other examples: light, waves,

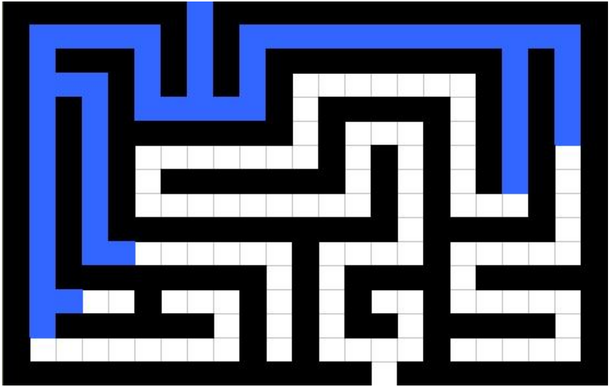
Example I: Flood Dynamics



Theorem: Every point is reached in minimal time.

Other examples: light, waves, viruses,

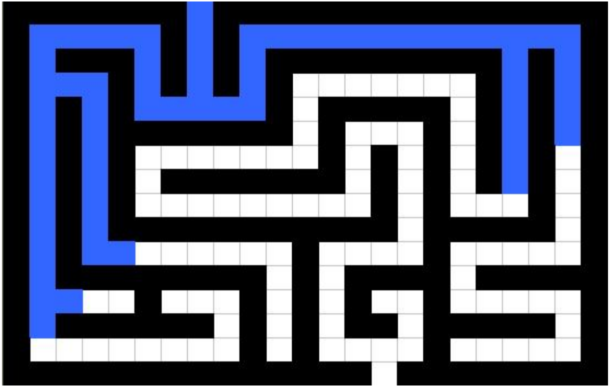
Example I: Flood Dynamics



Theorem: Every point is reached in minimal time.

Other examples: light, waves, viruses, rumors,

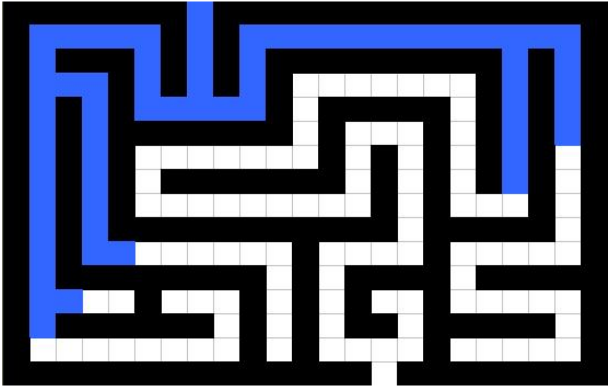
Example I: Flood Dynamics



Theorem: Every point is reached in minimal time.

Other examples: light, waves, viruses, rumors, etc.

Example I: Flood Dynamics

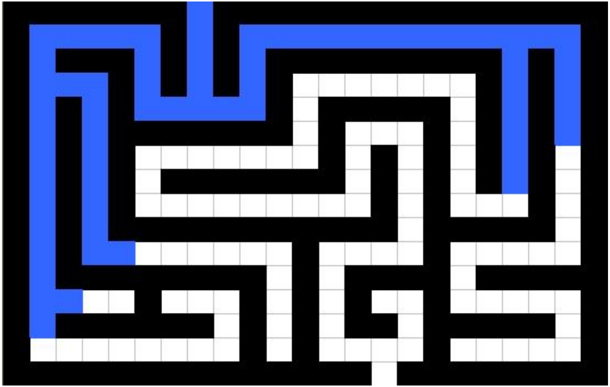


Theorem: Every point is reached in minimal time.

Other examples: light, waves, viruses, rumors, etc.

Algorithm:

Example I: Flood Dynamics

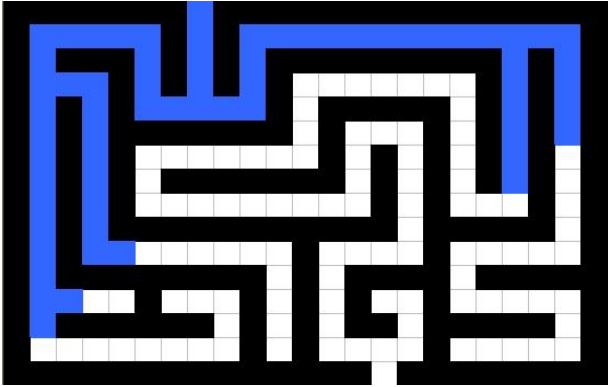


Theorem: Every point is reached in minimal time.

Other examples: light, waves, viruses, rumors, etc.

Algorithm: Shortest paths in the networks

Example I: Flood Dynamics

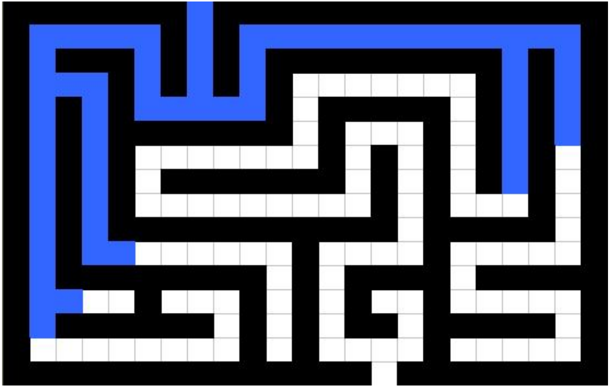


Theorem: Every point is reached in minimal time.

Other examples: light, waves, viruses, rumors, etc.

Algorithm: Shortest paths in the networks (Dijkstra, 1956).

Example I: Flood Dynamics



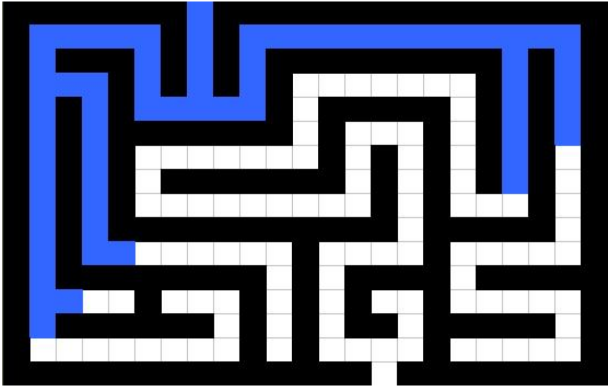
Theorem: Every point is reached in minimal time.

Other examples: light, waves, viruses, rumors, etc.

Algorithm: Shortest paths in the networks (Dijkstra, 1956).

NB:

Example I: Flood Dynamics



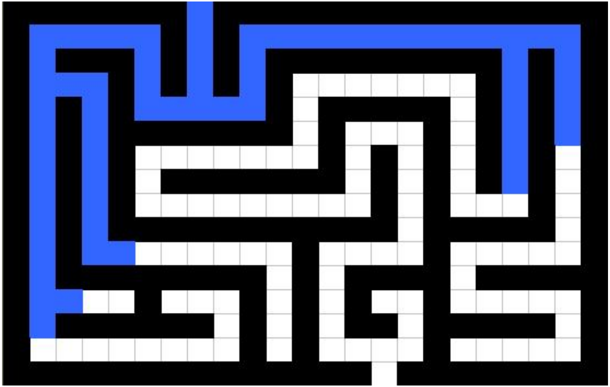
Theorem: Every point is reached in minimal time.

Other examples: light, waves, viruses, rumors, etc.

Algorithm: Shortest paths in the networks (Dijkstra, 1956).

NB: Continuous time

Example I: Flood Dynamics



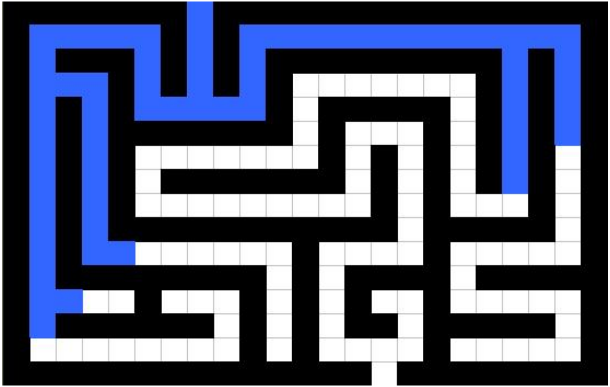
Theorem: Every point is reached in minimal time.

Other examples: light, waves, viruses, rumors, etc.

Algorithm: Shortest paths in the networks (Dijkstra, 1956).

NB: Continuous time $\Rightarrow$ Discrete time

Example I: Flood Dynamics



Theorem: Every point is reached in minimal time.

Other examples: light, waves, viruses, rumors, etc.

Algorithm: Shortest paths in the networks (Dijkstra, 1956).

NB: Continuous time $\Rightarrow$ Discrete time $\Rightarrow$ Polynomial complexity.

Example II: Superwised Growth

Example II: Superwised Growth

Model situation:

Example II: Superwised Growth

Model situation: Design of construction resisting to a load.

Example II: Superwised Growth

Model situation: Design of construction resisting to a load.

- We have a set elements of "thicknesses" x_i, $i = 1, \dots, n$.

Example II: Superwised Growth

Model situation: Design of construction resisting to a load.

- We have a set elements of "thicknesses" x_i, $i = 1, \ldots, n$.
- The elements "interact" at "nodes", unifying the set of "neighbors" $\sigma_j \subseteq \{1, \ldots, n\}$, $j = 1, \ldots, m$.

Example II: Superwised Growth

Model situation: Design of construction resisting to a load.

- We have a set elements of "thicknesses" x_i, $i = 1, \ldots, n$.
- The elements "interact" at "nodes", unifying the set of "neighbors" $\sigma_j \subseteq \{1, \ldots, n\}$, $j = 1, \ldots, m$.

Constraint:

Example II: Superwised Growth

Model situation: Design of construction resisting to a load.

- We have a set elements of "thicknesses" x_i, $i = 1, \ldots, n$.
- The elements "interact" at "nodes", unifying the set of "neighbors" $\sigma_j \subseteq \{1, \ldots, n\}$, $j = 1, \ldots, m$.

Constraint: The "tensions" at nodes are below certain limits.

Example II: Superwised Growth

Model situation: Design of construction resisting to a load.

- We have a set elements of "thicknesses" x_i, $i = 1, \ldots, n$.
- The elements "interact" at "nodes", unifying the set of "neighbors" $\sigma_j \subseteq \{1, \ldots, n\}$, $j = 1, \ldots, m$.

Constraint: The "tensions" at nodes are below certain limits.

Algorithm:

Example II: Superwised Growth

Model situation: Design of construction resisting to a load.

- We have a set elements of “thicknesses” x_i, $i = 1, \ldots, n$.
- The elements “interact” at “nodes”, unifying the set of ”neighbors” $\sigma_j \subseteq \{1, \ldots, n\}$, $j = 1, \ldots, m$.

Constraint: The “tensions” at nodes are below certain limits.

Algorithm: 1) Start from the zero load.

Example II: Superwised Growth

Model situation: Design of construction resisting to a load.

- We have a set elements of "thicknesses" x_i, $i = 1, \ldots, n$.
- The elements "interact" at "nodes", unifying the set of "neighbors" $\sigma_j \subseteq \{1, \ldots, n\}$, $j = 1, \ldots, m$.

Constraint: The "tensions" at nodes are below certain limits.

Algorithm: 1) Start from the zero load.

2) If all tensions are acceptable, increase the load.

Example II: Superwised Growth

Model situation: Design of construction resisting to a load.

- We have a set elements of "thicknesses" x_i, $i = 1, \dots, n$.
- The elements "interact" at "nodes", unifying the set of "neighbors" $\sigma_j \subseteq \{1, \dots, n\}$, $j = 1, \dots, m$.

Constraint: The "tensions" at nodes are below certain limits.

Algorithm: 1) Start from the zero load.

2) If all tensions are acceptable, increase the load.

3) Otherwise, enforce the neighbors at problematic nodes.

Example II: Superwised Growth

Model situation: Design of construction resisting to a load.

- We have a set elements of "thicknesses" x_i, $i = 1, \ldots, n$.
- The elements "interact" at "nodes", unifying the set of "neighbors" $\sigma_j \subseteq \{1, \ldots, n\}$, $j = 1, \ldots, m$.

Constraint: The "tensions" at nodes are below certain limits.

Algorithm: 1) Start from the zero load.

2) If all tensions are acceptable, increase the load.

3) Otherwise, enforce the neighbors at problematic nodes.

Interpretation:

Example II: Superwised Growth

Model situation: Design of construction resisting to a load.

- We have a set elements of "thicknesses" x_i, $i = 1, \dots, n$.
- The elements "interact" at "nodes", unifying the set of "neighbors" $\sigma_j \subseteq \{1, \dots, n\}$, $j = 1, \dots, m$.

Constraint: The "tensions" at nodes are below certain limits.

Algorithm: 1) Start from the zero load.

2) If all tensions are acceptable, increase the load.

3) Otherwise, enforce the neighbors at problematic nodes.

Interpretation: Switching subgradient method for solving the problem

Example II: Superwised Growth

Model situation: Design of construction resisting to a load.

- We have a set elements of "thicknesses" x_i, $i = 1, \ldots, n$.
- The elements "interact" at "nodes", unifying the set of "neighbors" $\sigma_j \subseteq \{1, \ldots, n\}$, $j = 1, \ldots, m$.

Constraint: The "tensions" at nodes are below certain limits.

Algorithm: 1) Start from the zero load.

2) If all tensions are acceptable, increase the load.

3) Otherwise, enforce the neighbors at problematic nodes.

Interpretation: Switching subgradient method for solving the problem

$$\max_{x \in X} \{ \langle f, x \rangle : \ f_j(x) \leq 0, \ j = 1, \ldots, m \}.$$

Example II: Superwised Growth

Model situation: Design of construction resisting to a load.

- We have a set elements of "thicknesses" x_i, $i = 1, \ldots, n$.
- The elements "interact" at "nodes", unifying the set of "neighbors" $\sigma_j \subseteq \{1, \ldots, n\}$, $j = 1, \ldots, m$.

Constraint: The "tensions" at nodes are below certain limits.

Algorithm: 1) Start from the zero load.

2) If all tensions are acceptable, increase the load.

3) Otherwise, enforce the neighbors at problematic nodes.

Interpretation: Switching subgradient method for solving the problem

$$\max_{x \in X} \{ \langle f, x \rangle : \; f_j(x) \leq 0, \; j = 1, \ldots, m \}.$$

Complexity:

Example II: Superwised Growth

Model situation: Design of construction resisting to a load.

- We have a set elements of "thicknesses" x_i, $i = 1, \ldots, n$.
- The elements "interact" at "nodes", unifying the set of "neighbors" $\sigma_j \subseteq \{1, \ldots, n\}$, $j = 1, \ldots, m$.

Constraint: The "tensions" at nodes are below certain limits.

Algorithm: 1) Start from the zero load.

2) If all tensions are acceptable, increase the load.

3) Otherwise, enforce the neighbors at problematic nodes.

Interpretation: Switching subgradient method for solving the problem

$$\max_{x \in X} \{ \langle f, x \rangle : \; f_j(x) \leq 0, \; j = 1, \ldots, m \}.$$

Complexity: $O(\frac{1}{\epsilon^2})$ iterations.

Example II: Superwised Growth

Model situation: Design of construction resisting to a load.

- We have a set elements of "thicknesses" x_i, $i = 1, \ldots, n$.
- The elements "interact" at "nodes", unifying the set of "neighbors" $\sigma_j \subseteq \{1, \ldots, n\}$, $j = 1, \ldots, m$.

Constraint: The "tensions" at nodes are below certain limits.

Algorithm: 1) Start from the zero load.

2) If all tensions are acceptable, increase the load.

3) Otherwise, enforce the neighbors at problematic nodes.

Interpretation: Switching subgradient method for solving the problem

$$\max_{x \in X} \{ \langle f, x \rangle : \ f_j(x) \leq 0, \ j = 1, \ldots, m \}.$$

Complexity: $O(\frac{1}{\epsilon^2})$ iterations.

NB: 1) Many naturally growing systems have hidden optimality.

Example II: Superwised Growth

Model situation: Design of construction resisting to a load.

- We have a set elements of "thicknesses" x_i, $i = 1, \ldots, n$.
- The elements "interact" at "nodes", unifying the set of "neighbors" $\sigma_j \subseteq \{1, \ldots, n\}$, $j = 1, \ldots, m$.

Constraint: The "tensions" at nodes are below certain limits.

Algorithm: 1) Start from the zero load.

2) If all tensions are acceptable, increase the load.

3) Otherwise, enforce the neighbors at problematic nodes.

Interpretation: Switching subgradient method for solving the problem

$$\max_{x \in X} \{ \langle f, x \rangle : \; f_j(x) \leq 0, \; j = 1, \ldots, m \}.$$

Complexity: $O(\frac{1}{\epsilon^2})$ iterations.

NB: 1) Many naturally growing systems have hidden optimality.

2) They are created by *unintentional* optimization.

Example III: Intuitive reaction

Example III: Intuitive reaction

Model situation:

Example III: Intuitive reaction

Model situation:

- Our current "position" is suddenly changed by a small random perturbation.

Example III: Intuitive reaction

Model situation:

- Our current “position” is suddenly changed by a small random perturbation.
- If we feel ourselves better, we do a big step in the same direction.

Example III: Intuitive reaction

Model situation:

- Our current "position" is suddenly changed by a small random perturbation.
- If we feel ourselves better, we do a big step in the same direction.
- If not, we do the big step in the opposite direction.

Example III: Intuitive reaction

Model situation:

- Our current “position” is suddenly changed by a small random perturbation.
- If we feel ourselves better, we do a big step in the same direction.
- If not, we do the big step in the opposite direction.

NB: Looks a bit unbalanced.

Example III: Intuitive reaction

Model situation:

- Our current "position" is suddenly changed by a small random perturbation.
- If we feel ourselves better, we do a big step in the same direction.
- If not, we do the big step in the opposite direction.

NB: Looks a bit unbalanced.

Interpretation: Random search for solving the $\min_{x \in \mathbb{R}^n} f(x)$:

Example III: Intuitive reaction

Model situation:

- Our current "position" is suddenly changed by a small random perturbation.
- If we feel ourselves better, we do a big step in the same direction.
- If not, we do the big step in the opposite direction.

NB: Looks a bit unbalanced.

Interpretation: Random search for solving the $\min_{x \in \mathbb{R}^n} f(x)$:

- Generate a Gaussian direction $u_k \in \mathbb{R}^n$.

Example III: Intuitive reaction

Model situation:

- Our current “position” is suddenly changed by a small random perturbation.
- If we feel ourselves better, we do a big step in the same direction.
- If not, we do the big step in the opposite direction.

NB: Looks a bit unbalanced.

Interpretation: Random search for solving the $\min_{x \in \mathbb{R}^n} f(x)$:

- Generate a Gaussian direction $u_k \in \mathbb{R}^n$.
- Update our position:

Example III: Intuitive reaction

Model situation:

- Our current "position" is suddenly changed by a small random perturbation.
- If we feel ourselves better, we do a big step in the same direction.
- If not, we do the big step in the opposite direction.

NB: Looks a bit unbalanced.

Interpretation: Random search for solving the $\min\limits_{x \in \mathbb{R}^n} f(x)$:

- Generate a Gaussian direction $u_k \in \mathbb{R}^n$.
- Update our position: $x_{k+1} = x_k - h_k \frac{f(x_k + \mu u_k) - f(x_k)}{\mu} u_k$.

Example III: Intuitive reaction

Model situation:

- Our current "position" is suddenly changed by a small random perturbation.
- If we feel ourselves better, we do a big step in the same direction.
- If not, we do the big step in the opposite direction.

NB: Looks a bit unbalanced.

Interpretation: Random search for solving the $\min_{x \in \mathbb{R}^n} f(x)$:

- Generate a Gaussian direction $u_k \in \mathbb{R}^n$.
- Update our position: $x_{k+1} = x_k - h_k \frac{f(x_k + \mu u_k) - f(x_k)}{\mu} u_k$.

NB: μ is very small,

Example III: Intuitive reaction

Model situation:

- Our current “position” is suddenly changed by a small random perturbation.
- If we feel ourselves better, we do a big step in the same direction.
- If not, we do the big step in the opposite direction.

NB: Looks a bit unbalanced.

Interpretation: Random search for solving the $\min_{x \in \mathbb{R}^n} f(x)$:

- Generate a Gaussian direction $u_k \in \mathbb{R}^n$.
- Update our position: $x_{k+1} = x_k - h_k \frac{f(x_k + \mu u_k) - f(x_k)}{\mu} u_k$.

NB: μ is very small, $h_k \approx \frac{1}{\sqrt{k}}$.

Example III: Intuitive reaction

Model situation:

- Our current "position" is suddenly changed by a small random perturbation.
- If we feel ourselves better, we do a big step in the same direction.
- If not, we do the big step in the opposite direction.

NB: Looks a bit unbalanced.

Interpretation: Random search for solving the $\min_{x \in \mathbb{R}^n} f(x)$:

- Generate a Gaussian direction $u_k \in \mathbb{R}^n$.
- Update our position: $x_{k+1} = x_k - h_k \frac{f(x_k + \mu u_k) - f(x_k)}{\mu} u_k$.

NB: μ is very small, $h_k \approx \frac{1}{\sqrt{k}}$.

Complexity: $O(\frac{n}{\epsilon^2})$ iterations.

Example III: Intuitive reaction

Model situation:

- Our current "position" is suddenly changed by a small random perturbation.
- If we feel ourselves better, we do a big step in the same direction.
- If not, we do the big step in the opposite direction.

NB: Looks a bit unbalanced.

Interpretation: Random search for solving the $\min\limits_{x \in \mathbb{R}^n} f(x)$:

- Generate a Gaussian direction $u_k \in \mathbb{R}^n$.
- Update our position: $x_{k+1} = x_k - h_k \frac{f(x_k + \mu u_k) - f(x_k)}{\mu} u_k$.

NB: μ is very small, $h_k \approx \frac{1}{\sqrt{k}}$.

Complexity: $O(\frac{n}{\epsilon^2})$ iterations.

Examples:

Example III: Intuitive reaction

Model situation:

- Our current "position" is suddenly changed by a small random perturbation.
- If we feel ourselves better, we do a big step in the same direction.
- If not, we do the big step in the opposite direction.

NB: Looks a bit unbalanced.

Interpretation: Random search for solving the $\min_{x \in \mathbb{R}^n} f(x)$:

- Generate a Gaussian direction $u_k \in \mathbb{R}^n$.
- Update our position: $x_{k+1} = x_k - h_k \frac{f(x_k + \mu u_k) - f(x_k)}{\mu} u_k$.

NB: μ is very small, $h_k \approx \frac{1}{\sqrt{k}}$.

Complexity: $O(\frac{n}{\epsilon^2})$ iterations.

Examples: Nature,

Example III: Intuitive reaction

Model situation:

- Our current "position" is suddenly changed by a small random perturbation.
- If we feel ourselves better, we do a big step in the same direction.
- If not, we do the big step in the opposite direction.

NB: Looks a bit unbalanced.

Interpretation: Random search for solving the $\min_{x \in \mathbb{R}^n} f(x)$:

- Generate a Gaussian direction $u_k \in \mathbb{R}^n$.
- Update our position: $x_{k+1} = x_k - h_k \frac{f(x_k + \mu u_k) - f(x_k)}{\mu} u_k$.

NB: μ is very small, $h_k \approx \frac{1}{\sqrt{k}}$.

Complexity: $O(\frac{n}{\epsilon^2})$ iterations.

Examples: Nature, Social life, etc.

Conclusion

Conclusion

... Many things to do ...

Conclusion

... Many things to do ...

Thank you for your attention!