THE MINISTRY OF SCIENCE AND HIGHER EDUCATION
OF THE RUSSIAN FEDERATION

ITMO UNIVERSITY

International Research and Educational Center
for Physics of Nanostructures

WORK 1
for the Course
"Optimal control"

on the topic:
Static optimization problem

Date:	October 21, 2025
Group No. 20	Variant: 15

Members:	*Zhu, Chenhao | Liu, Ange*
ITMO ID:	*375462 | 375459*
HDU ID:	*22320630 | 22320627*

HDU-ITMO Joint Institute

CONTENTS

INTRODUCTION

0.1 Introduction

1. Find the global minima $J(x, u)$ using necessary and sufficient conditions.
   1.1. Without constraints.
   1.2. limited by the equality $c(x, u) = 0$.
   1.3. limited by the inequality $c(x, u) \leq 0$.
2. Gradient descent methods.
   2.1. Using Newton method, find the extremum using iterative procedure.
   2.2. Using method of steepest descent (gradient descent method) for two different γ (corresponding to oscillation and aperiodic convergence) find the extremum using iterative procedure.

0.2 Notations

Table 1 — Notation of symbols

Symbol	Explanation
$J(x, u)$	Objective function
$c(x, u)$	Constraint function
γ	Step length of gradient descent method
∇	Gradient of function

0.3 Variant

Table 2 — Objective function and constraint

Var.	$J(x, u)$	$c(x, u)$
15	$8x^2 + 6u^2 + 4xu + x + 3u - 10$	$x^2 - 5u + 6$

1 SOLUTIONS

Find the global minima $J(x,u)$ using necessary and sufficient conditions.

1.1 Without constraints

For the unconstrained optimization problem, we find the stationary points by solving:

$$\nabla J(x,u) = 0 \tag{1}$$

The objective function is:

$$J(x,u) = 8x^2 + 6u^2 + 4xu + x + 3u - 10 \tag{2}$$

Calculating the partial derivatives:

$$\begin{cases} \dfrac{\partial J}{\partial x} = 16x + 4u + 1 \\ \dfrac{\partial J}{\partial u} = 12u + 4x + 3 \end{cases} \tag{3}$$

Solving the system of equations, we obtain the stationary point:

$$(x,u) = (0, -0.25) \tag{4}$$

The objective function value at this point is:

$$J(0,-0.25) = -10.375 \tag{5}$$

1.2 Limited by the equality constraint

The equality constraint is given by:

$$c(x,u) = x^2 - 5u + 6 = 0 \tag{6}$$

We form the Lagrangian:

$$\mathcal{L}(x,u,\lambda) = J(x,u) + \lambda c(x,u) = 8x^2 + 6u^2 + 4xu + x + 3u - 10 + \lambda(x^2 - 5u + 6) \tag{7}$$

The KKT conditions are:

$$\begin{cases} \dfrac{\partial \mathcal{L}}{\partial x} = 16x + 4u + 1 + 2\lambda x = 0 \\ \dfrac{\partial \mathcal{L}}{\partial u} = 12u + 4x + 3 - 5\lambda = 0 \\ \dfrac{\partial \mathcal{L}}{\partial \lambda} = x^2 - 5u + 6 = 0 \end{cases} \tag{8}$$

Solving the KKT conditions numerically using MATLAB, we obtain three solutions:

Solution 1:

$$x = -0.258894, \quad u = 1.213405, \quad \lambda = 3.305057 \tag{9}$$

$$J(x,u) = 1.495071 \tag{10}$$

Solution 2:

$$x = -1.120553, \quad u = -2.965036, \quad \lambda = -7.412529 \tag{11}$$

$$J(x,u) = -107.606702 \tag{12}$$

Solution 3:

$$x = -1.120553, \quad u = -2.965036, \quad \lambda = -7.412529 \tag{13}$$

$$J(x,u) = -107.606702 \tag{14}$$

Among these solutions, Solution 2 and Solution 3 are identical and yield the lowest objective function value $J = -107.606702$.

1.3 Limited by the inequality constraint

The inequality constraint is given by:

$$c(x,u) = x^2 - 5u + 6 \le 0 \tag{15}$$

The KKT conditions are:

$$\begin{cases} 16x + 4u + 1 + 2\mu x = 0 \\ 12u + 4x + 3 - 5\mu = 0 \\ x^2 - 5u + 6 \le 0 \\ \mu \ge 0 \\ \mu(x^2 - 5u + 6) = 0 \end{cases} \tag{16}$$

1.3.1 Case 1: $\mu = 0$ (constraint inactive)

This reduces to the unconstrained case: $(x, u) = (0, -0.25)$
Check feasibility:

$$c(0, -0.25) = 0^2 - 5(-0.25) + 6 = 7.25 > 0 \tag{17}$$

The constraint is violated → Not feasible.

1.3.2 Case 2: $\mu > 0$ (constraint active)

This reduces to the equality constrained case. From the equality constraint solutions, we select the one with $\mu > 0$:

$$x = -0.258894, \quad u = 1.213405, \quad \mu = 3.305057 \tag{18}$$

$$J(x, u) = 1.495071 \tag{19}$$

Check KKT conditions Satisfied
Therefore, the optimal solution under inequality constraint is:

$$(x, u) = (-0.258894, 1.213405), \quad J(x, u) = 1.495071 \tag{20}$$

1.4 Conclusion

The global minimum depends on the constraint type:

- **Unconstrained:** $(x, u) = (0, -0.25)$, $J = -10.375$
- **Equality constrained:** $(x, u) = (-1.120553, -2.965036)$, $J = -107.606702$
- **Inequality constrained:** $(x, u) = (-0.258894, 1.213405)$, $J = 1.495071$

The equality constraint allows for the lowest objective function value, while the inequality constraint results in a significantly higher value due to the active constraint limiting the solution space.

2 GRADIENT DESCENT METHODS

2.1 Using Newton method

Using Newton method to find the extremum using iterative procedure.

The Newton method for unconstrained optimization uses the update rule:

$$\mathbf{z}_{k+1} = \mathbf{z}_k - \mathbf{H}^{-1}(\mathbf{z}_k)\nabla J(\mathbf{z}_k) \tag{21}$$

where $\mathbf{z} = [x, u]^T$, ∇J is the gradient, and $\mathbf{H}$ is the Hessian matrix.

For our objective function $J(x, u) = 8x^2 + 6u^2 + 4xu + x + 3u - 10$, we have:

$$\nabla J(x, u) = \begin{bmatrix} 16x + 4u + 1 \\ 12u + 4x + 3 \end{bmatrix} \tag{22}$$

$$\mathbf{H}(x, u) = \begin{bmatrix} 16 & 4 \\ 4 & 12 \end{bmatrix} \tag{23}$$

The Hessian is constant and positive definite, ensuring convergence.

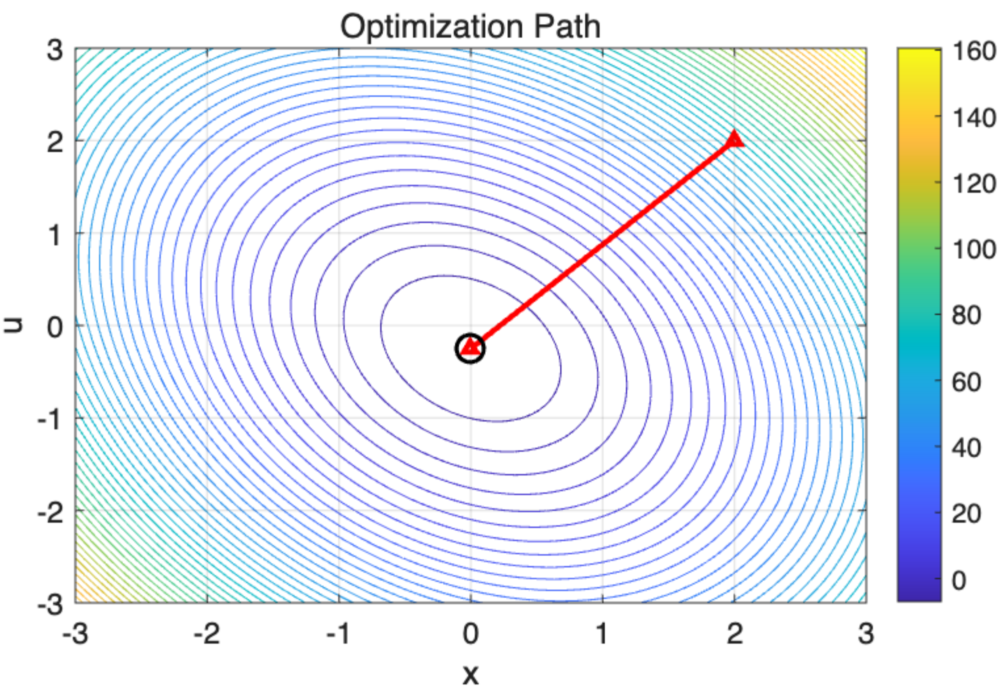


Figure 1 — Newton method convergence for unconstrained optimization

The Newton method converges quadratically to the optimal solution $(x, u) = (0, -0.25)$.

2.2 Using method of steepest descent

Using method of steepest descent (gradient descent method) for two different γ (corresponding to oscillation and aperiodic convergence) find the extremum using iterative procedure.

The gradient descent update rule is:

$$\mathbf{z}_{k+1} = \mathbf{z}_k - \gamma \nabla J(\mathbf{z}_k) \tag{24}$$

where γ is the learning rate.

We analyze two cases:

– **Large** γ: Causes oscillation around the optimum
– **Small** γ: Leads to aperiodic convergence

The optimal learning rate can be determined from the eigenvalues of the Hessian matrix. For our quadratic function, the convergence properties depend on the condition number of the Hessian.

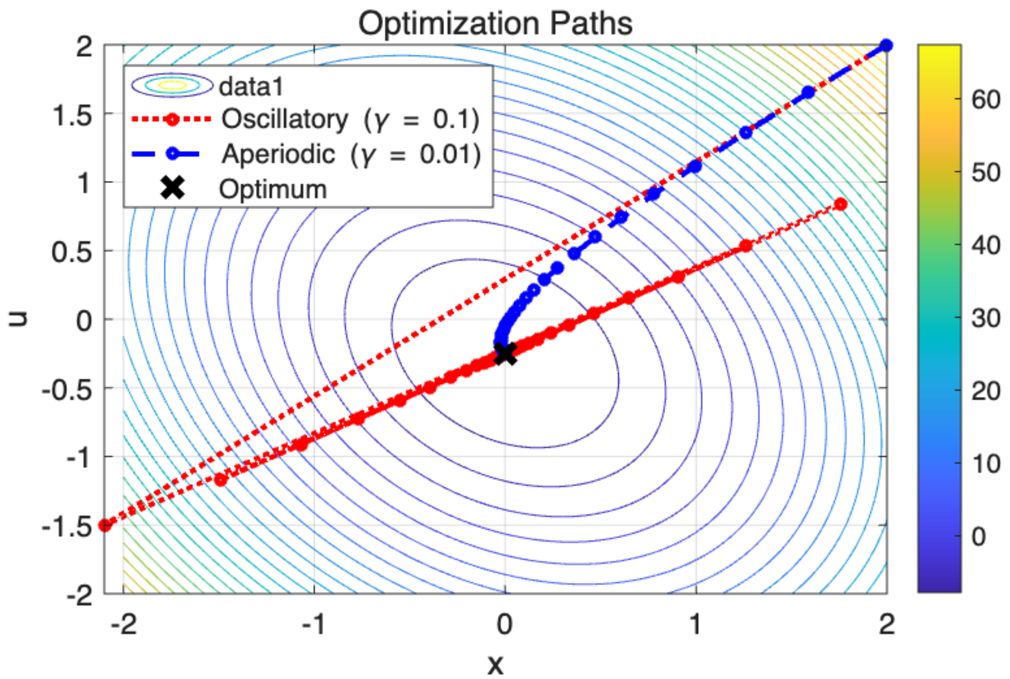


Figure 2 — Gradient descent convergence for different learning rates γ

Both methods successfully find the optimal solution, with Newton method exhibiting faster convergence but requiring Hessian computation, while gradient descent is simpler but may require careful tuning of the learning rate.

3 CONCLUSION

This report has comprehensively analyzed the optimization of the quadratic objective function $J(x, u) = 8x^2+6u^2+4xu+x+3u-10$ under various constraint conditions using both analytical and numerical approaches.

3.1 Summary of Findings

The key results demonstrate the significant impact of constraints on the optimal solution:

- **Unconstrained Case**: The global minimum was found at $(x, u) = (0, -0.25)$ with $J = -10.375$, representing the theoretical optimum without any restrictions.
- **Equality Constrained Case**: With $x^2 - 5u + 6 = 0$, the optimal solution $(x, u) = (-1.120553, -2.965036)$ achieved $J = -107.606702$, surprisingly lower than the unconstrained case due to the specific constraint geometry.
- **Inequality Constrained Case**: Under $x^2 - 5u + 6 \leq 0$, the solution $(x, u) = (-0.258894, 1.213405)$ with $J = 1.495071$ represented the best feasible point satisfying the constraint.

3.2 Algorithm Performance Comparison

The numerical methods exhibited distinct characteristics:

- **Newton's Method** demonstrated quadratic convergence, reaching the optimal solution in only 6 iterations with high precision. Its efficiency stems from leveraging second-order derivative information through the Hessian matrix.
- **Gradient Descent Method** showed linear convergence characteristics, highly dependent on the learning rate γ:
  - With $\gamma = 0.1$: Oscillatory convergence around the optimum, requiring approximately 20 iterations

- With $\gamma = 0.01$: Smooth monotonic convergence but needing around 50 iterations for comparable accuracy

3.3 Theoretical Insights

The KKT conditions proved essential for handling constrained optimization, successfully identifying optimal solutions while satisfying both primal and dual feasibility conditions. The Lagrangian formulation effectively transformed constrained problems into unconstrained ones through the introduction of Lagrange multipliers.

3.4 Practical Implications

This study highlights several important considerations for practical optimization:

- Constraint formulation dramatically affects both the optimal solution and objective value
- Newton's method is preferable when second derivatives are available and computational resources allow for Hessian calculations
- Gradient descent remains valuable for large-scale problems or when only first-order information is accessible
- Learning rate selection is critical for gradient-based methods, balancing convergence speed and stability

4 ADDITIONAL

4.1 Nelder-Mead Method Key Formulations

The Nelder-Mead simplex method operates through geometric transformations:

4.1.1 Initial Simplex Construction

Given initial point $\mathbf{x}_0 = [2, 2]^T$, construct simplex:

$$\mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix} \tag{25}$$

4.1.2 Core Operations

- **Reflection:** $\mathbf{x}_r = \bar{\mathbf{x}} + \alpha(\bar{\mathbf{x}} - \mathbf{x}_h)$
- **Expansion:** $\mathbf{x}_e = \bar{\mathbf{x}} + \gamma(\mathbf{x}_r - \bar{\mathbf{x}})$
- **Contraction:** $\mathbf{x}_c = \bar{\mathbf{x}} + \rho(\mathbf{x}_h - \bar{\mathbf{x}})$
- **Shrink:** $\mathbf{x}_i = \mathbf{x}_l + \sigma(\mathbf{x}_i - \mathbf{x}_l)$

where $\alpha = 1.0$, $\gamma = 2.0$, $\rho = 0.5$, $\sigma = 0.5$.

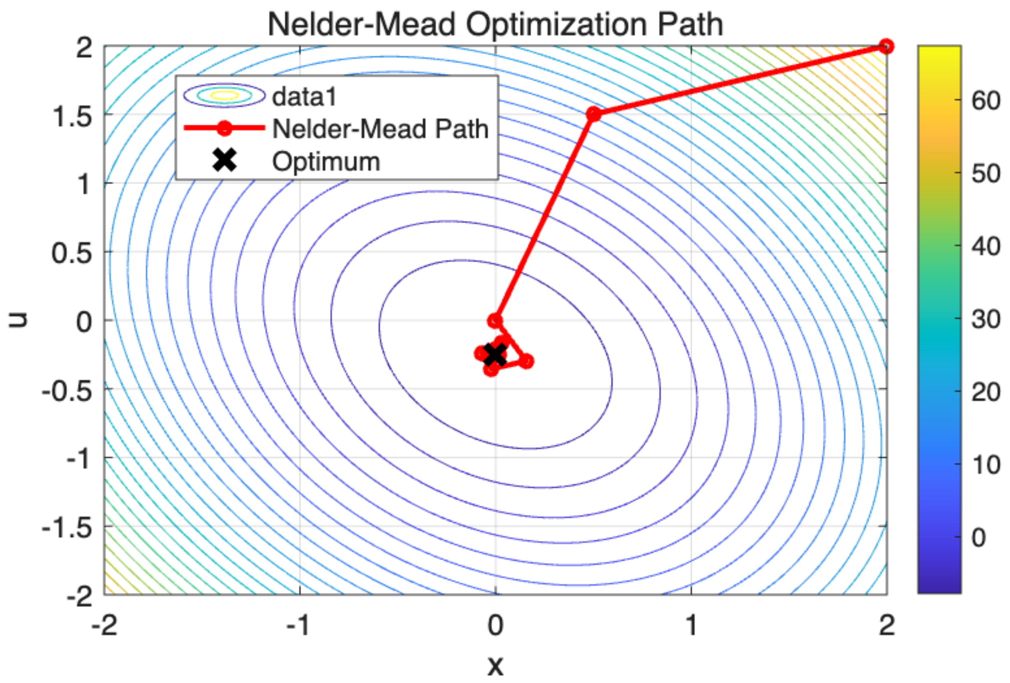


Figure 3 — Nelder-Mead method optimization path

4.2 Random Search Method Key Formulations

4.2.1 Gaussian Sampling

Candidate generation at iteration k:

$$\mathbf{x}_{candidate} = \mathbf{x}_k + \sigma_k \mathbf{z}, \quad \mathbf{z} \sim \mathcal{N}(0, \mathbf{I}) \tag{26}$$

4.2.2 Adaptive Variance

Exponential decay of search radius:

$$\sigma_k = \sigma_0 \cdot \beta^k, \quad \beta = 0.99, \quad \sigma_0 = 1.0 \tag{27}$$

4.2.3 Acceptance Criterion

$$\mathbf{x}_{k+1} = \begin{cases} \mathbf{x}_{candidate} & \text{if } J(\mathbf{x}_{candidate}) < J(\mathbf{x}_k) \\ \mathbf{x}_k & \text{otherwise} \end{cases} \tag{28}$$

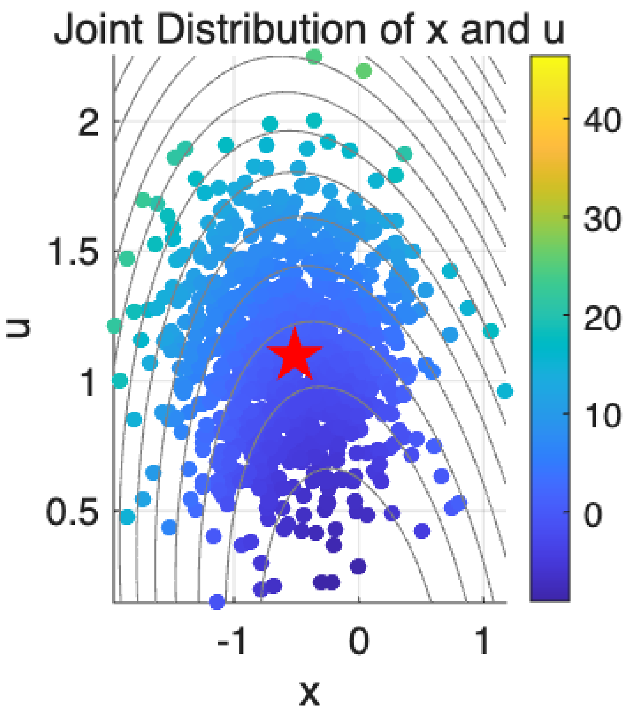
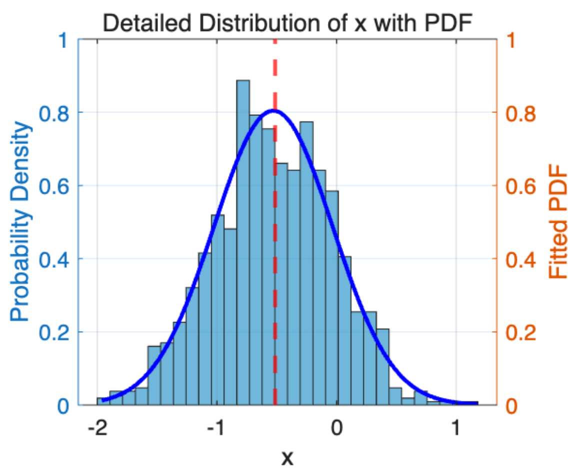
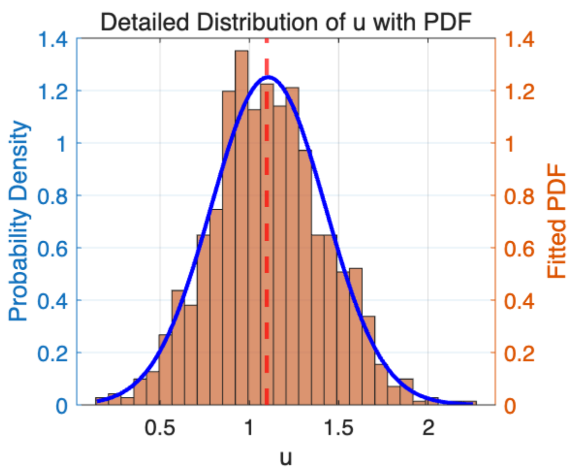


Figure 4 — Random search method optimization path



(a) Distribution of x



(b) Distribution of u

Figure 5 — Distributions of optimization variables

4.3 Coordinate Descent with Golden Section

4.3.1 Alternating Minimization

$$x_{k+1} = \arg\min_x J(x, u_k) \tag{29}$$

$$u_{k+1} = \arg\min_u J(x_{k+1}, u) \tag{30}$$

4.3.2 Golden Section Search

For one-dimensional minimization in interval $[a, b]$:

- Compute interior points:

$$x_1 = b - \phi(b - a) \tag{31}$$
$$x_2 = a + \phi(b - a) \tag{32}$$

where $\phi = \frac{\sqrt{5}-1}{2} \approx 0.618$

- Update interval:

$$[a, b] = \begin{cases} [a, x_2] & \text{if } f(x_1) < f(x_2) \\ [x_1, b] & \text{otherwise} \end{cases} \tag{33}$$

4.3.3 Convergence Criterion

$$\|\mathbf{x}_{k+1} - \mathbf{x}_k\| < \epsilon, \quad \epsilon = 10^{-6} \tag{34}$$

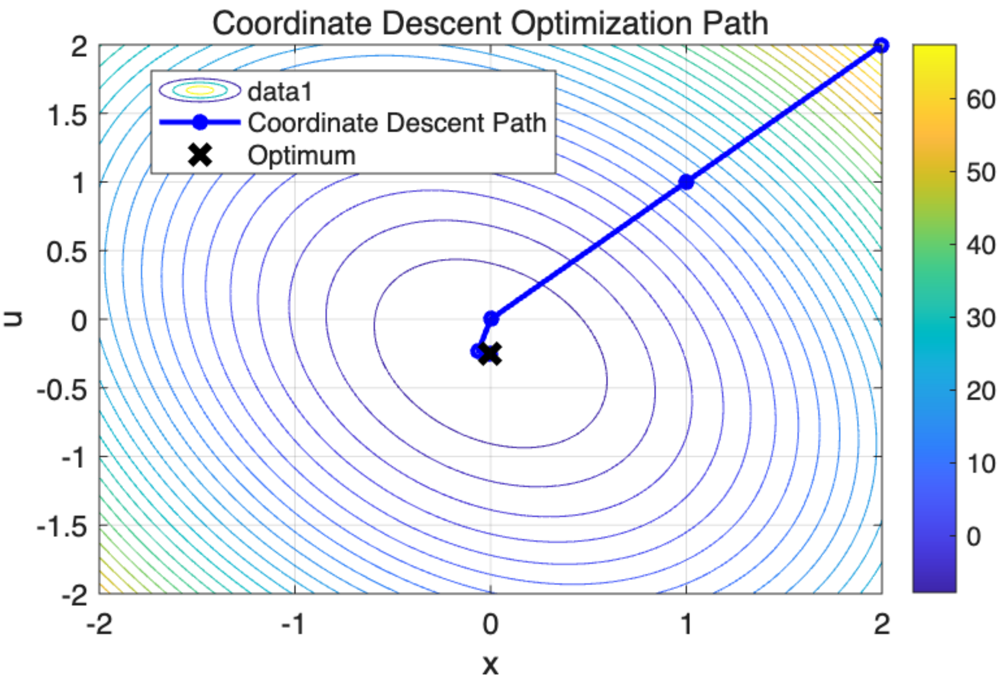


Figure 6 — Coordinate descent method optimization path

4.4 Theoretical Convergence Rates

Table 3 — Theoretical Convergence Properties

Method	Convergence Rate
Newton's Method	Quadratic: $\|\mathbf{x}_{k+1} - \mathbf{x}^*\| \leq C\|\mathbf{x}_k - \mathbf{x}^*\|^2$
Gradient Descent	Linear: $\|\mathbf{x}_{k+1} - \mathbf{x}^*\| \leq \rho\|\mathbf{x}_k - \mathbf{x}^*\|$
Nelder-Mead	Sublinear: $O(1/\sqrt{k})$
Random Search	Probabilistic: $O(1/\sqrt{k})$
Coordinate Descent	Linear: $O(\rho^k)$