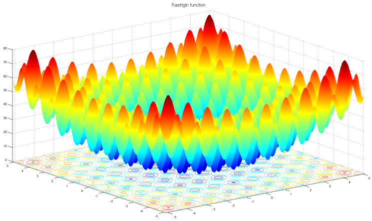
Quantum-Inspired Gradient Descent in Finance

Winston Bartle

July 12, 2024



Importance in Finance: Portfolio Optimization

Classic Gradient Descent:

- Used in traditional methods like Mean-Variance Optimization
- Balances risk and return
- Can get stuck in local optima, leading to suboptimal portfolios

Quantum-Inspired Gradient Descent:

- Emulates quantum behavior for probabilistic exploration
- Escape local optima more effectively
- Lead to more robust portfolio allocations
- Better navigation of global risk-return landscape

Rastrigin Function

- A common test function for optimization algorithms
- Highly non-convex with many local minima
- Formula:

$$f(x, y) = 20 + x^2 + y^2 - 10(\cos(2\pi x) + \cos(2\pi y))$$

- Global minimum at (0,0)

Gradient Descent

- Basic optimization algorithm
- Update rule:

$$x_{n+1} = x_n - \alpha \nabla f(x_n)$$

  where α is the learning rate
- Can get stuck in local minima

Stochastic Gradient Descent

- Variant of gradient descent for large-scale problems
- Uses random subset of data in each iteration
- Update rule:

$$x_{n+1} = x_n - \alpha \nabla f_i(x_n)$$

  where f_i is the cost function for a random subset i
- Limited advantages:
  - Faster iterations, but not necessarily faster convergence
  - May escape some shallow local minima due to noise
  - Lower memory requirements, but at cost of accuracy
- Significant challenges:
  - Noisy updates lead to unstable convergence
  - Requires careful learning rate scheduling
  - Still prone to getting trapped in deeper local minima
  - Less effective for complex, non-convex landscapes

Quantum-Inspired Gradient Descent

- Inspired by quantum superposition and measurement principles
- Multiple "walkers" explore the space simultaneously
- Periodic "collapse" to best solution
- Update rule (for each walker):

$$x_{n+1} = x_n - \alpha \nabla f(x_n) + \epsilon$$

  where ϵ is quantum-inspired noise
- Key advantages:
  - Superior exploration of complex, non-convex landscapes
  - Higher probability of finding global optimum
  - Inherent ability to escape deep local minima
  - Maintains solution diversity throughout optimization
  - Potentially faster convergence for multi-modal problems

Quantum-Inspired Noise in QIGD

Examples of quantum-inspired noise:

- Gaussian noise with quantum amplitude:

$$\epsilon = A \cdot \mathcal{N}(0, 1) \cdot \exp(-t/T)$$

  where A is amplitude, t is current iteration, T is total iterations

- Quantum tunneling-inspired noise:

$$\epsilon = B \cdot \text{sign}(\nabla f(x)) \cdot \exp(-C|\nabla f(x)|)$$

  where B and C are tunable parameters

Quantum-Inspired Noise in QIGD (Con'd)

- Superposition-inspired noise:

$$\epsilon = D \cdot (\sin(\omega t) + \cos(\omega t))$$

  where D and ω are amplitude and frequency parameters

- Entanglement-inspired noise:

$$\epsilon_i = E \cdot \sum_{j \neq i} (x_j - x_i)$$

  where E is a coupling strength, i and j are walker indices

Simulation Setup

- Function: 2D Rastrigin
- Algorithms compared: Classic GD, SGD, QIGDs
- Number of iterations: 100
- Learning rate: 0.0001
- QIGD: 5 walkers
- Starting point: Random within [-5, 5] range

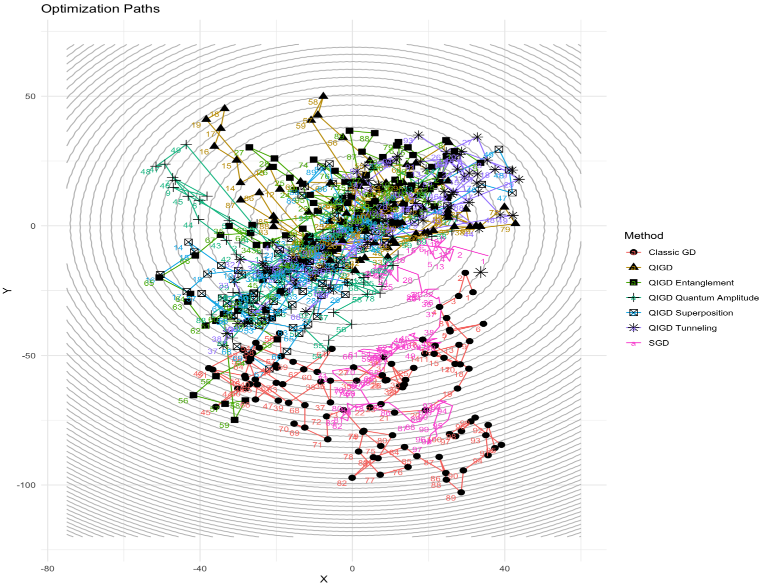
Key Assumptions

- No boundary constraints imposed
- Quantum-inspired "collapse" every 10 iterations
- Performance measured by final position of walkers

Results and Interpretation

- Global minimum and value: (0,0) & 0
- Classic GD final position and value: (25.52, -80.33) & 7114.316
- Stochastic GD final position and value: (22.9 -70.0) & 5444.758
- QIGDs final positions and value:
  - Classic Gaussian: (-8.288916, -8.395365) & 159.066
  - Quantum Amplitude: (-1.694835, -1.783685) & 21.43029
  - Tunneling: (0.3847438, 0.387603) & 9.328606
  - Superposition: (5.493589, 5.45238) & 60.267
  - Entanglement: (-6.412945, -6.314809) & 97.18657

Visualization and Analysis



Conclusions and Future Work

- QIGD shows promise in finding global minimum
- Challenges with noise selection
- Future work:
  - Fine-tune parameters
  - Explore alternative quantum-inspired techniques