**Numerical Optimization Homework #5**

**20161216 HAN SANGJUN**

장세현, 원경호 조교님

**Problem 1)** Find local minimum of f(x, y) = 2\*x^2 + 5\*y^2

- The global minimum of f(x, y) is 0 when (x, y) = (0, 0)

**Problem 2)** Find local minimum of

f(x, y) = (1.5 - x + x\*y)^2 + (2.25 - x + x\*y^2)^2 + (2.625 - x + x\*y^3)^2

- The global minimum of f(x, y) is 0 when (x, y) = (3, 0.5)

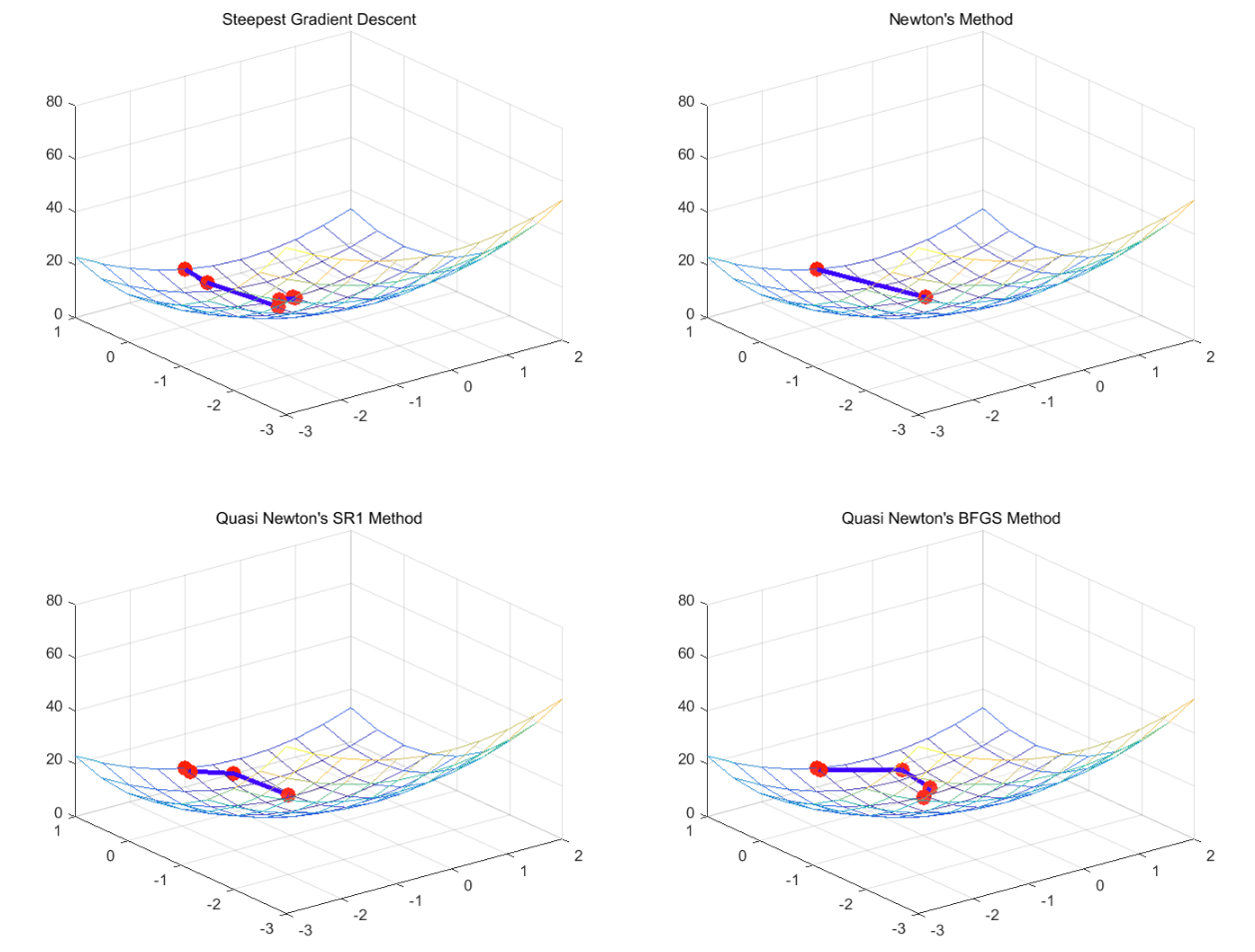
**Problem 3)** Find local minimum of f(x, y) = 100 \* (y - x^2)^2 + 3 \* (1 - x)^2

- The global minimum of f(x, y) is 0 when (x, y) = (1, 1)

**1. When initial point (x, y) = (-1, 1)**

Problem 1) Find local minimum of f(x, y) = 2\*x^2 + 5\*y^2

- Compare gradient descent, Newton’s method, Quasi-SR1 and Quasi-BFGS



- Performance

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Method | Optimal value | Time (sec) | Iterations | Time / Iterations |
| Gradient descent | 0.000165 | 0.906 | 8 | 0.113 |
| Newton’s method | 0 | 0.057 | 2 | 0.029 |
| Qusai-SR1 | 0.021014 | 0.406 | 4 | 0.101 |
| Qusai-BFGS | 0.002289 | 0.649 | 6 | 0.108 |

- Discussion

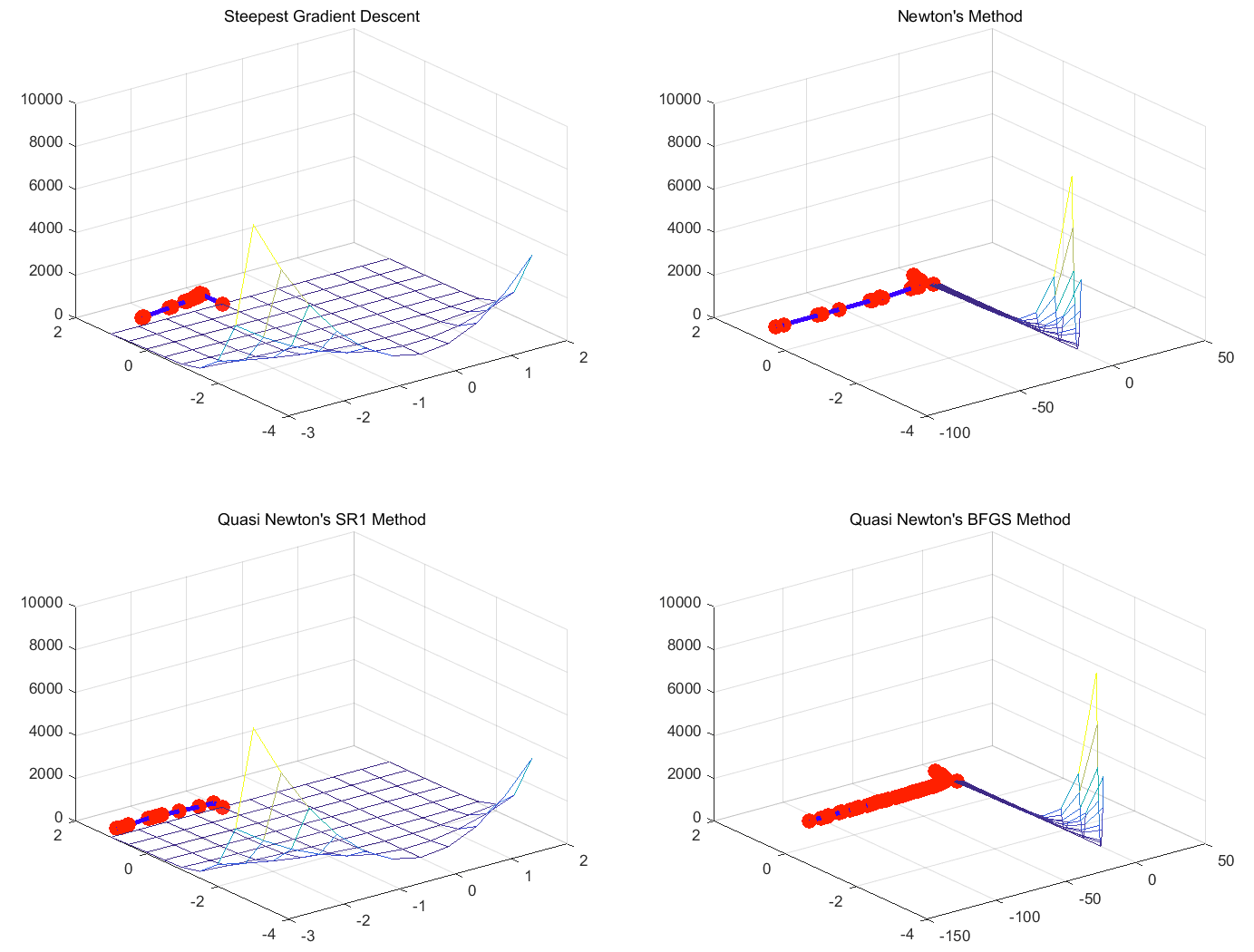
a. In terms of time and iteration, gradient descent took longest time and needed many iterations

b. It seems that if Hessian matrix can be computed fast, Newton’s method outperforms other method

Problem 2) Find local minimum of

f(x, y) = (1.5 - x + x\*y)^2 + (2.25 - x + x\*y^2)^2 + (2.625 - x + x\*y^3)^2

- Compare gradient descent, Newton’s method, Quasi-SR1 and Quasi-BFGS



- Performance

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Method | Optimal value | Time (sec) | Iterations | Time / Iterations |
| Gradient descent | 1.025 | 8.056 | 58 | 0.139 |
| Newton’s method | 0.470 | 3.436 | 30 | 0.115 |
| Qusai-SR1 | 0.941 | 1.738 | 14 | 0.124 |
| Qusai-BFGS | 0.467 | 4.571 | 73 | 0.063 |

- Discussion

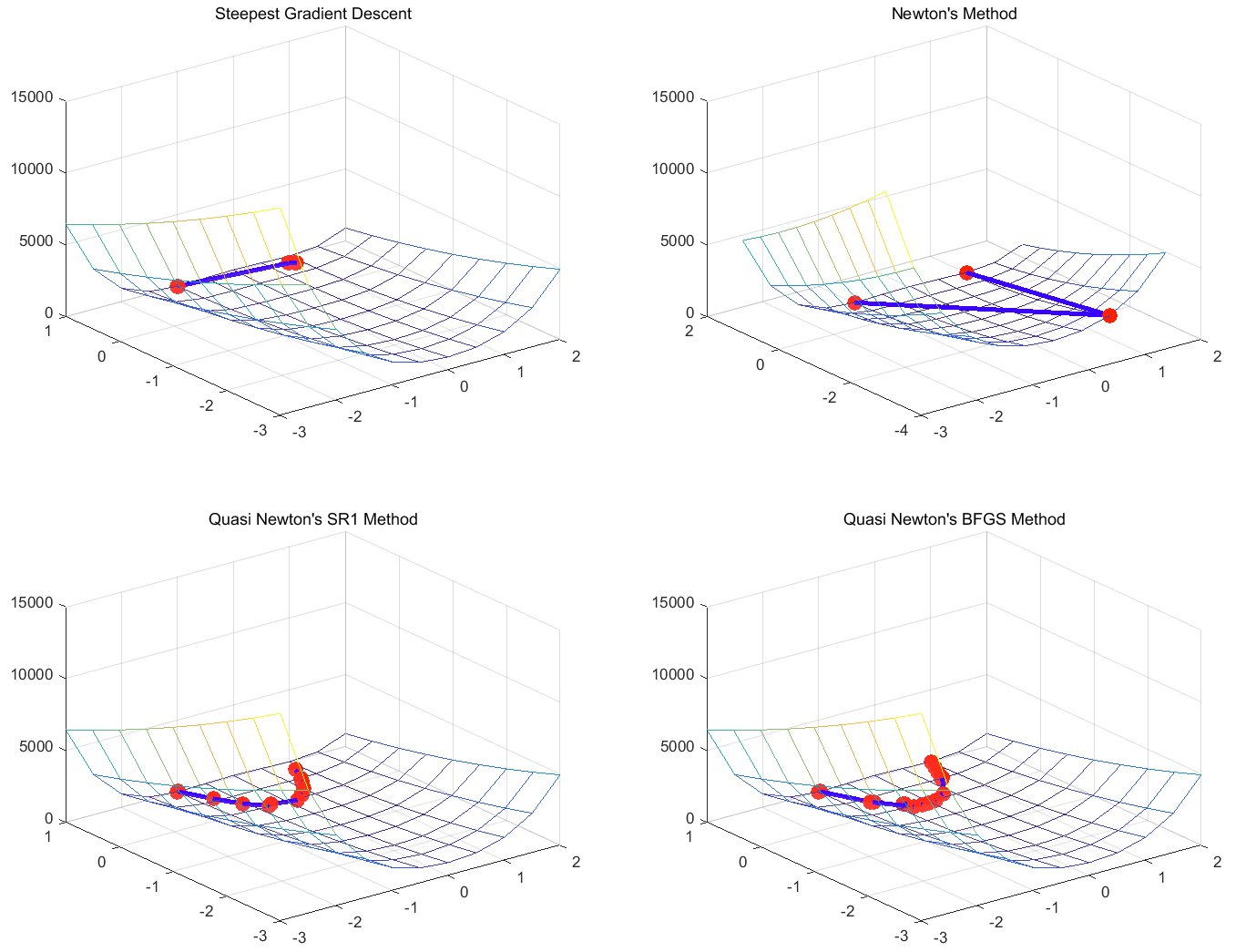
a. This function is so complex to find global minimum

b. Gradient descent and Qusai-SR1 converged to similar local minimum, but Qusai-SR1 was fast in terms of time and time/iterations

c. Newton’s method and Qusai-BFGS converged to other point, but they have same optimal value

Problem 3) Find local minimum of Find local minimum of f(x, y) = 100 \* (y - x^2)^2 + 3 \* (1 - x)^2

- Compare gradient descent, Newton’s method, Quasi-SR1 and Quasi-BFGS



- Performance

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Method | Optimal value | Time (sec) | Iterations | Time / Iterations |
| Gradient descent | 0.039 | 1.233 | 7 | 0.176 |
| Newton’s method | 0 | 0.079 | 3 | 0.026 |
| Qusai-SR1 | 0.060 | 1.403 | 17 | 0.083 |
| Qusai-BFGS | 0 | 1.433 | 19 | 0.075 |

- Discussion

a. Gradient descent converged slowly for one iterations (in terms of time/iteration)

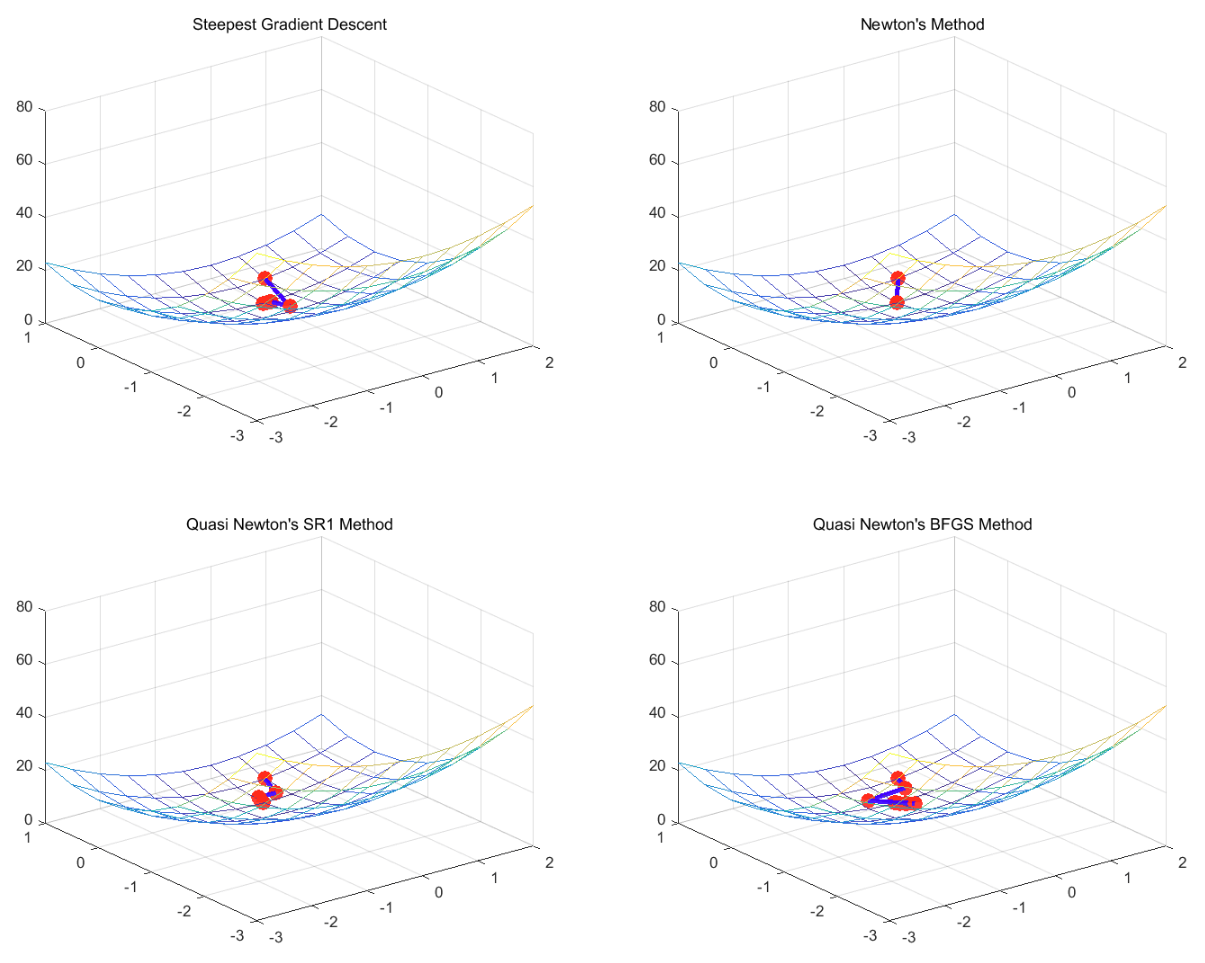
b. In the Qusai-Newton’s method, at some initial iterations, they converged to wrong direction because the Hessian matrix was not well approximated

c. Newton method outperforms other method if Hessian matrix was easily computed

**2. Initial point near global minimum**

Problem 1) Find local minimum of f(x, y) = 2\*x^2 + 5\*y^2 (initial point = (0.5, 0.5))

- Compare gradient descent, Newton’s method, Quasi-SR1 and Quasi-BFGS



- Performance

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Method | Optimal value | Time (sec) | Iterations | Time / Iterations |
| Gradient descent | 0 | 0.698 | 7 | 0.010 |
| Newton’s method | 0 | 0.052 | 2 | 0.026 |
| Qusai-SR1 | 0 | 0.404 | 6 | 0.067 |
| Qusai-BFGS | 0 | 0.537 | 9 | 0.060 |

- Discussion

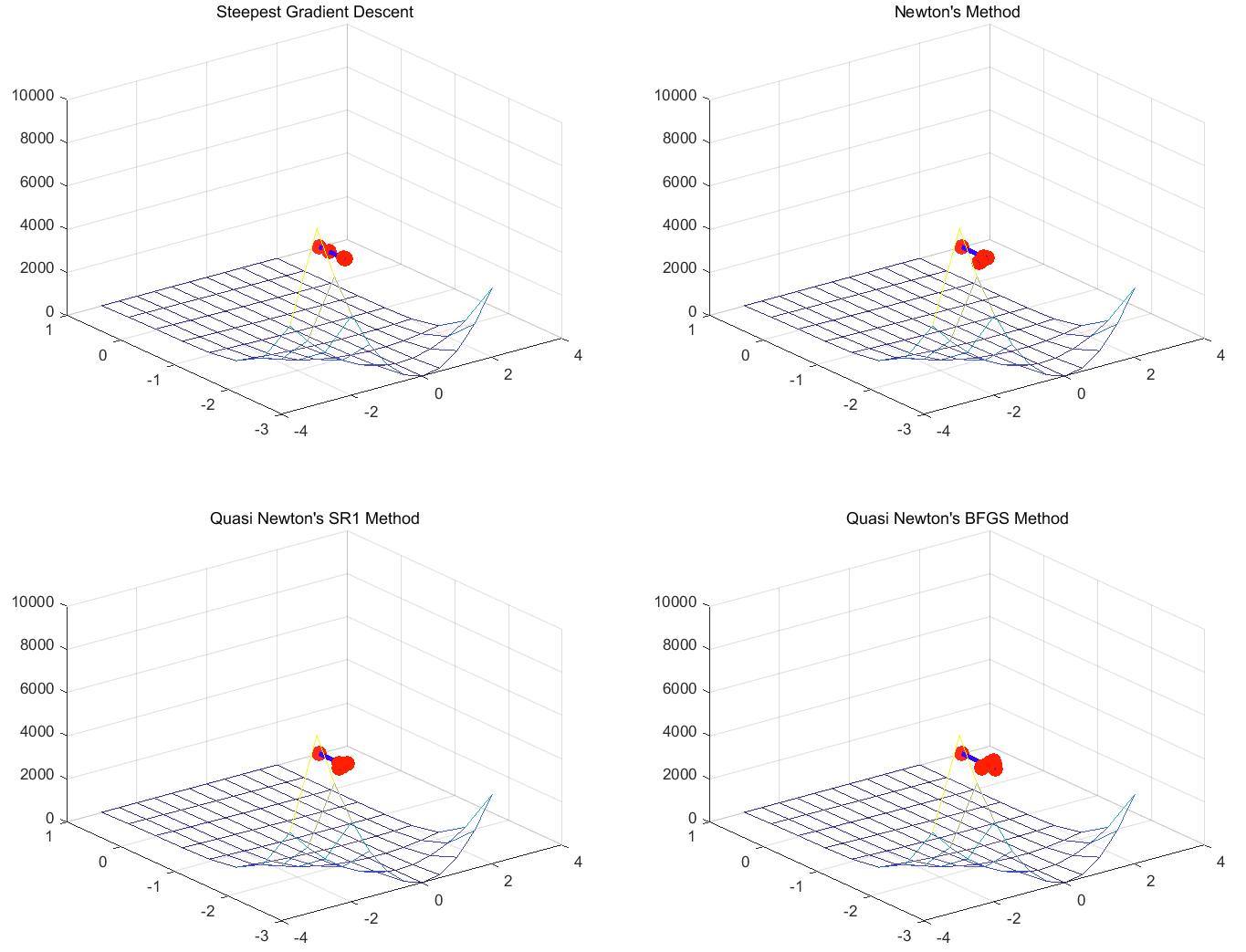
a. Newton’s method was fastest and gradient descent was slowest in terms of time

b. It is hard to evaluate performance in this case, because they are so close

Problem 2) Find local minimum of (initial point = (3.2, 1))

f(x, y) = (1.5 - x + x\*y)^2 + (2.25 - x + x\*y^2)^2 + (2.625 - x + x\*y^3)^2

- Compare gradient descent, Newton’s method, Quasi-SR1 and Quasi-BFGS



- Performance

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Method | Optimal value | Time (sec) | Iterations | Time / Iterations |
| Gradient descent | 0 | 1.042 | 6 | 0.174 |
| Newton’s method | 0 | 0.307 | 5 | 0.061 |
| Qusai-SR1 | 0 | 0.448 | 6 | 0.747 |
| Qusai-BFGS | 0 | 1.074 | 14 | 0.077 |

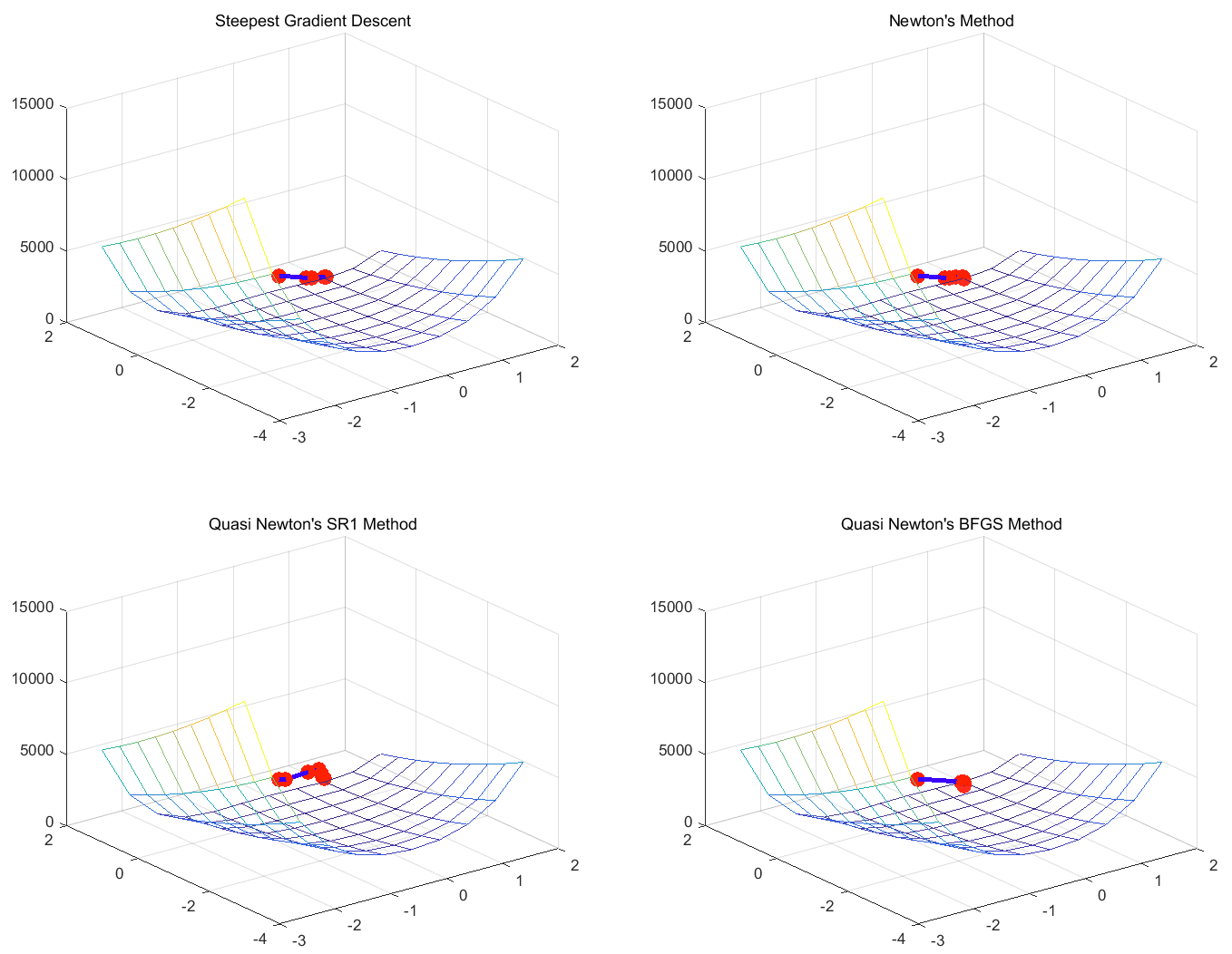
- Discussion

a. Newton’s method was fastest and Qusai-BFGS was slowest in terms of time

b. It is hard to evaluate performance in this case, because they are so close

Problem 3) Find local minimum of Find local minimum of f(x, y) = 100 \* (y - x^2)^2 + 3 \* (1 - x)^2

- Compare gradient descent, Newton’s method, Quasi-SR1 and Quasi-BFGS (initial point = (0.5, 1.5))



- Performance

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Method | Optimal value | Time (sec) | Iterations | Time / Iterations |
| Gradient descent | 0.003 | 0.939 | 5 | 0.188 |
| Newton’s method | 0 | 1.029 | 7 | 0.147 |
| Qusai-SR1 | 0.017 | 1.283 | 8 | 0.160 |
| Qusai-BFGS | 0 | 0.393 | 5 | 0.079 |

- Discussion

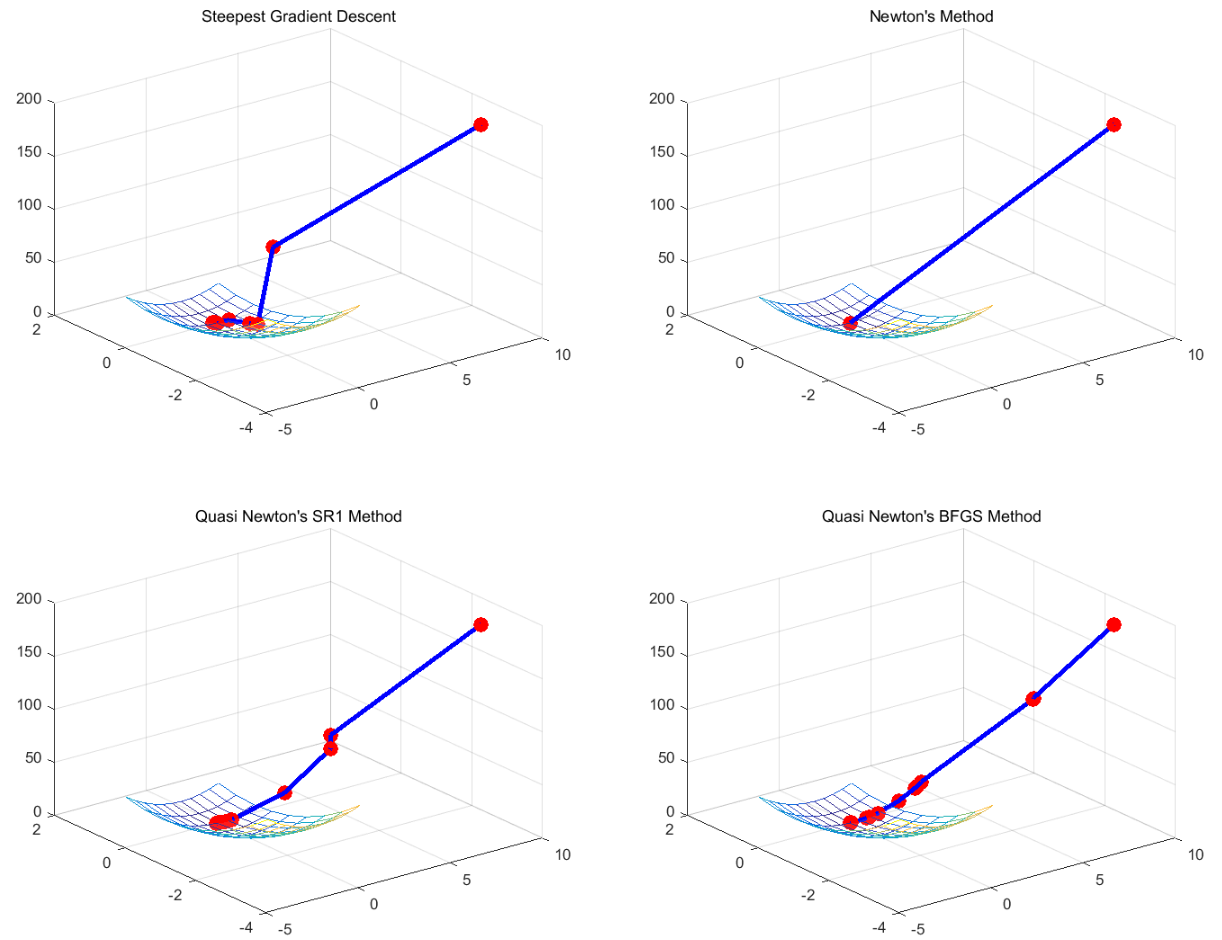
a. Qusai-BFGS was fastest and Qusai-SR1 was slowest in terms of time

b. It is hard to evaluate performance in this case, because they are so close

**3. Random initial point**

Problem 1) Find local minimum of f(x, y) = 2\*x^2 + 5\*y^2 (initial point = (8.6, -3.0))

- Compare gradient descent, Newton’s method, Quasi-SR1 and Quasi-BFGS



- Performance

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Method | Optimal value | Time (sec) | Iterations | Time / Iterations |
| Gradient descent | 0 | 0.907 | 10 | 0.090 |
| Newton’s method | 0 | 0.052 | 2 | 0.026 |
| Qusai-SR1 | 0 | 0.576 | 9 | 0.064 |
| Qusai-BFGS | 0 | 0.897 | 12 | 0.074 |

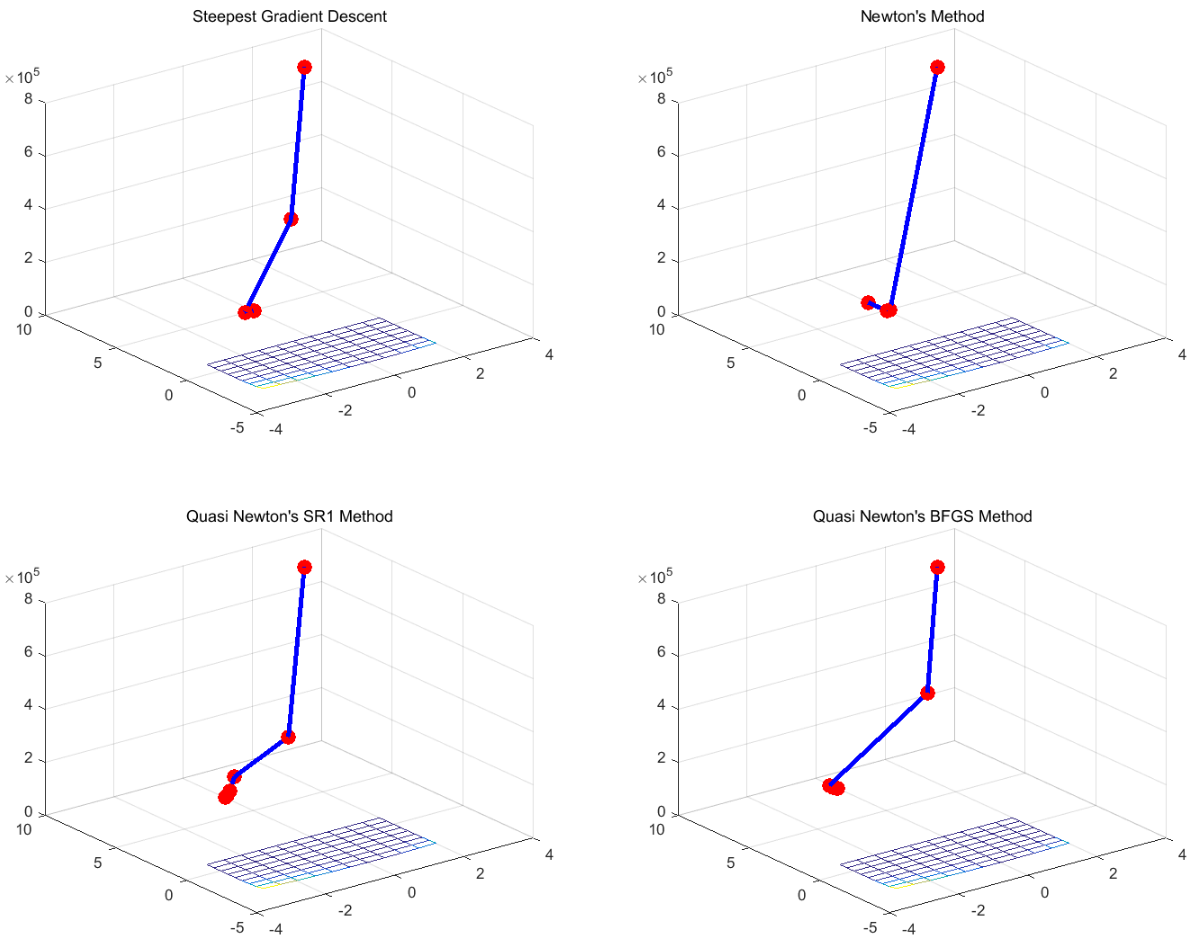
- Discussion

a. Newton’s method was fastest and gradient descent was slowest in terms of time and time/iterations

Problem 2) Find local minimum of (initial point = (2.3, 7.1))

f(x, y) = (1.5 - x + x\*y)^2 + (2.25 - x + x\*y^2)^2 + (2.625 - x + x\*y^3)^2

- Compare gradient descent, Newton’s method, Quasi-SR1 and Quasi-BFGS



- Performance

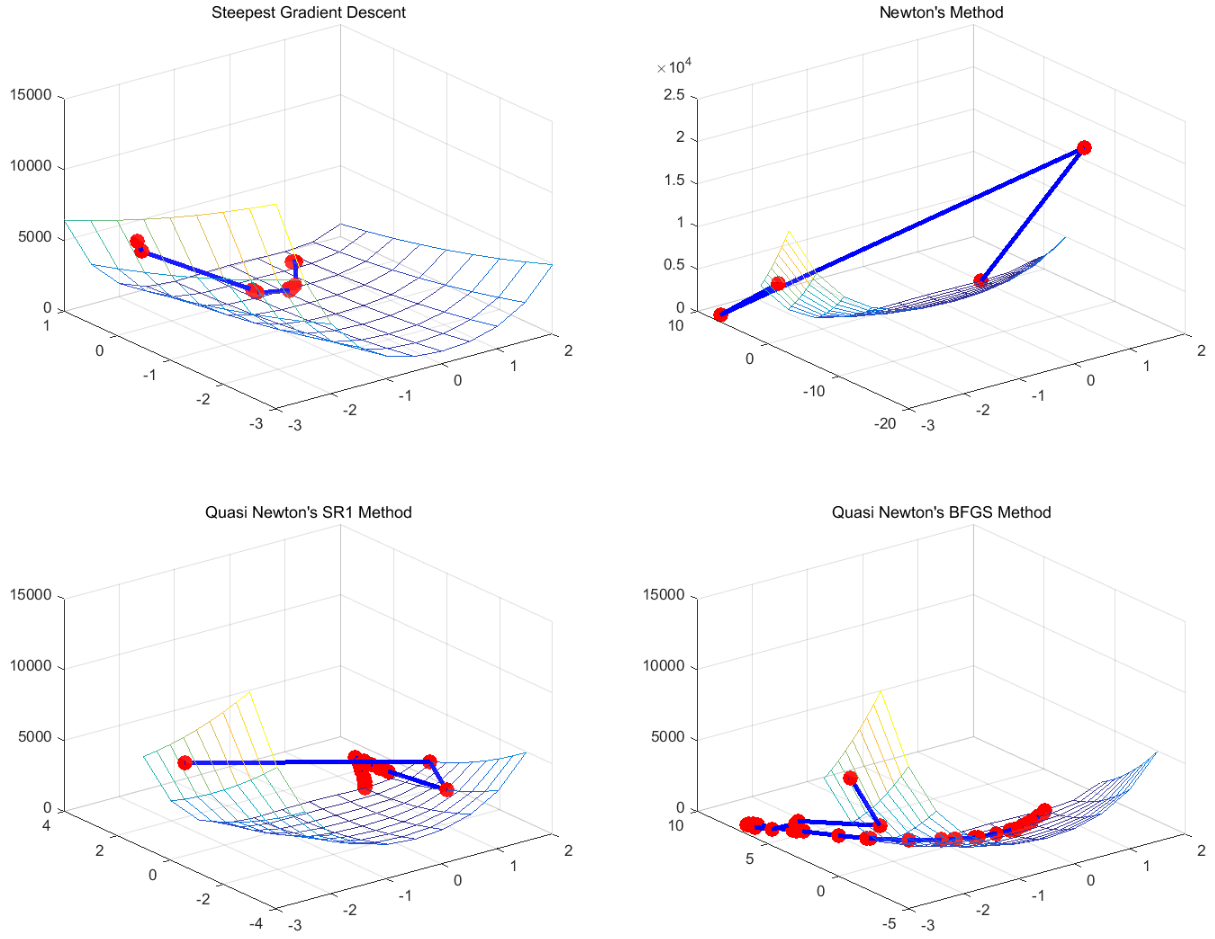
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Method | Optimal value | Time (sec) | Iterations | Time / Iterations |
| Gradient descent | 4.893 | 2.456 | 10 | 0.246 |
| Newton’s method | 5.376 | 2.163 | 9 | 0.240 |
| Qusai-SR1 | 5.577 | 1.170 | 13 | 0.090 |
| Qusai-BFGS | 5.878 | 1.228 | 12 | 0.102 |

- Discussion

a. Qusai-SR1 was fastest and gradient descent was slowest in terms of time and time/iterations, It is inferred that starting Hessian matrix is well approximated in Qusai-SR1

Problem 3) Find local minimum of Find local minimum of f(x, y) = 100 \* (y - x^2)^2 + 3 \* (1 - x)^2

- Compare gradient descent, Newton’s method, Quasi-SR1 and Quasi-BFGS (initial point = (-2.8, -0.2))



- Performance

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Method | Optimal value | Time (sec) | Iterations | Time / Iterations |
| Gradient descent | 0.090 | 2.113 | 14 | 0.152 |
| Newton’s method | 0 | 0.099 | 4 | 0.025 |
| Qusai-SR1 | 0 | 1.778 | 23 | 0.008 |
| Qusai-BFGS | 0 | 2.694 | 40 | 0.067 |

- Discussion

a. Newton’s method was fastest and gradient descent was slowest time/iterations, and Qusai-BFGS needed many iterations