

Worked solutions

Convexity and unconstrained optimization

Use these steps after attempting the worksheet. Numerical values are rounded only for display; the notebook uses full floating-point precision.

1. Convex sets

Classify $C1 = \{x: x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 4\}$ and $C2 = \{x: x_1^2 + x_2^2 \geq 1\}$. Prove convexity or give two feasible endpoints with an infeasible midpoint.

Step 1. $C1$ is the intersection of three half-spaces. Each inequality is affine, so the intersection is convex.

Step 2. For $C2$, choose $p = (-1, 0)$ and $q = (1, 0)$. Both have squared norm 1 and are feasible.

Step 3. The midpoint $(p+q)/2 = (0, 0)$ has squared norm $0 < 1$. This single counterexample proves that $C2$ is nonconvex.

2. Function classification

On $\mathbb{R}$, classify x^4 , $-x^2$, and $q(x) = (x^2 - 1)^2$. State whether a global minimizer exists and whether it is unique.

Step 1. For x^4 , $f'' = 12x^2 \geq 0$ everywhere, so f is convex. Its derivative $4x^3$ is strictly increasing, so it is strictly convex.

Step 2. $x^4 \geq 0$, with equality only at $x = 0$. Hence its global minimum is uniquely 0. It is not strongly convex on $\mathbb{R}$: $f''(0) = 0$ rules out a uniform positive lower curvature bound.

Step 3. For $-x^2$, $f'' = -2 < 0$. It is concave, and $f(x)$ tends to $-\infty$ as $|x|$ grows. There is no global minimum. The stationary point $x = 0$ is a global maximum.

Step 4. For q , $q(-1) = q(1) = 0$ while $q(0) = 1$. The midpoint inequality fails, proving nonconvexity. Because q is a square, it is nonnegative and has exactly two global minimizers, -1 and 1 .

3. Two gradient descent steps

Let $f(u, v) = (u - 1)^2 + 4(v + 2)^2$. Start at $x_0 = (3, 0)$. Use $\alpha = 0.1$ to compute two gradient descent steps, including the gradient and objective at each iterate.

Step 1. Gradient: $g = (2(u - 1), 8(v + 2))$. Hessian: $\text{diag}(2, 8)$. At x_0 , $f_0 = 20$ and $g_0 = (4, 16)$.

Step 2. $x_1 = x_0 - 0.1g_0 = (3 - 0.4, 0 - 1.6) = (2.6, -1.6)$.

Step 3. $f_1 = (1.6)^2 + 4(0.4)^2 = 2.56 + 0.64 = 3.2$. Recompute $g_1 = (3.2, 3.2)$.

Step 4. $x_2 = x_1 - 0.1g_1 = (2.28, -1.92)$. Then $f_2 = (1.28)^2 + 4(0.08)^2 = 1.6384 + 0.0256 = 1.664$.

Step 5. $g_2 = (2.56, 0.64)$ is not zero. The iterates improved the objective but have not reached the exact solution $(1, -2)$.

4. The step-size boundary

For the same quadratic, derive the range of constant positive α that converges from every initial point. Explain $\alpha=0.25$ and $\alpha=0.30$ from $x_0=(3,0)$.

Step 1. Write errors $e_u=u-1$ and $e_v=v+2$. Updating by $-\alpha \cdot g$ gives $e_{u,new}=(1-2\alpha)e_u$ and $e_{v,new}=(1-8\alpha)e_v$.

Step 2. A geometric sequence converges to zero for every initial error if the absolute multiplier is below 1.

Step 3. $|1-2\alpha|<1$ gives $0<\alpha<1$; $|1-8\alpha|<1$ gives $0<\alpha<1/4$. Their intersection is $0<\alpha<1/4$.

Step 4. At $\alpha=1/4$, $e_{u,k}=2(1/2)^k$ and $e_{v,k}=2(-1)^k$. Thus $f_k=4(1/4)^k+16$ tends to 16, not 0. The v coordinate alternates between 0 and -4 .

Step 5. At $\alpha=0.30$, the v -error multiplier is -1.4 . Its magnitude grows from the nonzero initial error. The first iterate is $(1.8, -4.8)$, with $f=32>20$.

5. Newton and a coupled quadratic (Practice A)

Minimize $h(x,y)=x^2+2xy+3y^2-4x-8y+7$ on $\mathbb{R}^2$. Derive the gradient, Hessian, eigenvalues, stationary point, and its classification.

Step 1. Gradient: $(2x+2y-4, 2x+6y-8)$. Hessian rows: $(2,2)$ and $(2,6)$.

Step 2. The characteristic polynomial is $(2-\lambda)(6-\lambda)-4 = \lambda^2-8\lambda+8$.

Step 3. The roots are $[8\pm\sqrt{64-32}]/2 = 4\pm2\sqrt{2}$, both positive. The constant Hessian is positive definite everywhere.

Step 4. Stationarity gives $x+y=2$ and $x+3y=4$. Subtraction yields $2y=2$, so $y=1$ and $x=1$.

Step 5. $h(1,1)=1+2+3-4-8+7=1$. The stationary point is the unique global minimum.

Step 6. Independent check: writing $e=x-1$ and $d=y-1$ gives $h=1+(e+d)^2+2d^2\geq 1$, with equality only when $e=d=0$.

6. Nonquadratic Newton (Practice B)

Take one full Newton step for $r(x)=x^4+x^2$ starting at $x_0=2$. Check the new objective. Has the method reached the minimizer?

Step 1. $r'(x)=4x^3+2x$; $r''(x)=12x^2+2$. At $x_0=2$, $r'=36$ and $r''=50$.

Step 2. Solve $50p=-36$, giving $p=-0.72$. With a full step, $x_1=2-0.72=1.28$.

Step 3. $r(2)=20$; $r(1.28)=1.28^4+1.28^2=2.68435456+1.6384=4.32275456$.

Step 4. The objective decreases, but $r'(1.28)=10.948608$ is not zero. The exact minimizer is $x=0$ because r is nonnegative and equals zero only there.

Step 5. Unlike a positive-definite quadratic, this nonquadratic function is not generally solved by one Newton step.

7. Stationarity and a failed Newton step

For $q(x)=(x^2-1)^2$, calculate a full Newton step from 0.2. Explain why it increases q . Separately explain why x^4 and $-x^4$ cannot be classified from their first two derivatives at zero (Practice C).

Step 1. $q'=4x(x^2-1)$ and $q''=12x^2-4$. At 0.2, $g=-0.768$ and $H=-3.52$.

Step 2. $p=-g/H=-12/55\approx-0.2181818$. Therefore $x_1=1/5-12/55=-1/55\approx-0.0181818$.

Step 3. $g^*p\approx0.1675636>0$: this Newton direction is an ascent direction. Direct evaluation confirms $q(0.2)=0.9216$ and $q(x_1)=(3024/3025)^2\approx0.99933895$.

Step 4. A negative Hessian makes the local quadratic model concave; its stationary point is a model maximum, not a minimum.

Step 5. At zero, x^4 and $-x^4$ both have zero first and second derivatives. Yet $x^4\geq0$ gives a strict minimum, while $-x^4\leq0$ gives a strict maximum. A zero Hessian is inconclusive.

8. Lab interpretation

Run the five fixed step sizes and the five nonconvex starts in the notebook. Report the iterate, objective, gradient norm, stopping reason, and your classification. Do not use solver success alone as proof.

Step 1. For the quadratic, $\alpha=0.05, 0.10, 0.24$ converge toward $(1, -2)$. The $\alpha=0.24$ run oscillates in v while its magnitude contracts; $\alpha=0.25$ does not converge to the minimizer; $\alpha=0.30$ grows.

Step 2. For q with $\alpha=0.05$, starts -1.5 and -0.2 approach -1 ; starts 0.2 and 1.5 approach 1 . At these minima, $q=0$ and $q''=8>0$.

Step 3. Starting exactly at zero stops immediately under a gradient tolerance, but $q''(0)=-4<0$ identifies a strict local maximum.

Step 4. The notebook includes exact-value checks and central-difference checks. Compare numerical tolerances, not exact floating-point equality.

Closing questions: answers

A positive Hessian at one point does not prove global convexity. The Hessian criterion must hold everywhere on the open convex domain.

At $\alpha=0.25$ the v error alternates without shrinking. For the specified start, the objective tends to 16.

Newton increases q from 0.2 because the negative curvature produces $g^*p>0$. A stationary point of the local model need not minimize that model.