PS 8

Due: Fri, May 1

1: Nearest Points For a given point $(a, b) \in \mathbb{R}^2$, we want to find the nearest point (x, y) that lies on the hyperbola $xy = 1$.

1. Write a Lagrangian function $L(x, y, \lambda)$ such that the desired point is a stationary point of L.
2. Write a Newton iteration to find the stationary point of L for $(a, b) = (3, 4)$. Use the starting guess $(x, y, \lambda) = (a, b, 0)$, and demonstrate quadratic convergence.

Note: As the point is to demonstrate a knowledge of Lagrange multipliers, we will not give credit for solutions that eliminate the constraint in advance.

Answer: Add the Lagrange multiplier to the distance objective:

$$L(x, y, \lambda) = (x - a)^2/2 + (y - b)^2/2 + \lambda(xy - 1)$$

For Newton's iteration, we need the gradient

$$\nabla L = \begin{bmatrix} (x-a) + \lambda y \\ (y-b) + \lambda x \\ xy - 1 \end{bmatrix}$$

and the Hessian

$$H = \begin{bmatrix} 1 & \lambda & y \\ \lambda & 1 & x \\ y & x & 0 \end{bmatrix}$$

Starting from the initial guess $(x, y, \lambda) = (a, b, 0)$, we have the convergence history shown in Figure 1 to $(x, y, \lambda) = (0.26236, 3.8116, 0.71825)$.

2: Nonlinear Least Squares In this problem, we consider a nonlinear least squares fitting problem in which we fit the coefficient vector β defining a rational function

$$f(x; \beta) = \frac{\beta_1 + \beta_2 x + \beta_3 x^2 + \beta_4 x^3}{1 + \beta_5 x + \beta_6 x^2 + \beta_7 x^3}$$

by minimizing

$$\phi(\beta) = \sum_j (f(x_j; \beta) - y_j)^2$$

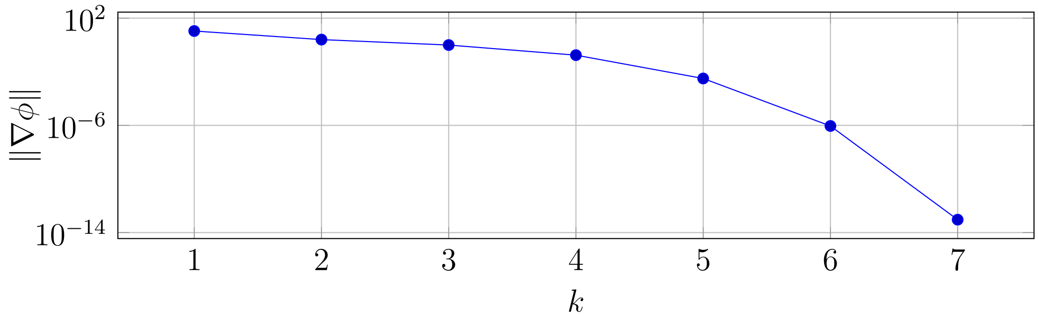

Figure 1: Convergence of Newton's iteration for projection onto a hyperbola.

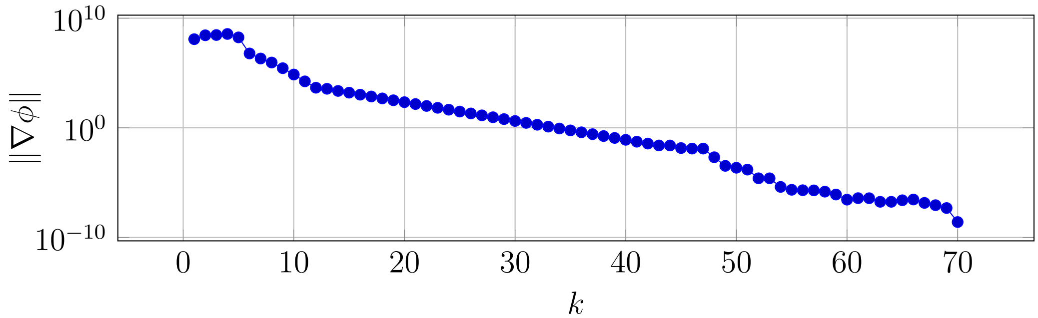

Figure 2: Convergence of Gauss-Newton for Thurber example

Your task: Complete the MATLAB script `ps8thuber.m` by filling in the code marked `TODO` with an appropriate solver iteration. You may use Gauss-Newton or Levenberg-Marquardt; I used Gauss-Newton with a line search (necessary to achieve convergence). Terminate when $\|J^T(f-y)\| < 10^{-8}$.

Answer: I used a Gauss-Newton iteration with line search; see the full `ps8thurber.m` for details. The convergence behavior, shown in Figure 2, is somewhat more complicated than in other examples we have done. In part, this is because we frequently require a reduced step size in order to guarantee progress.

3: Descent directions Suppose that H is symmetric and positive definite, and let $\tilde{p}$ be approximate $-H^{-1}\nabla\phi$, with residual

$$r = H\tilde{p} + \nabla\phi(x).$$

If $\kappa(H) = \lambda_{\max}(H)/\lambda_{\min}(H)$, show that if $\kappa(H)\|r\| < \|\nabla\phi\|$ then $\tilde{p}$ is a descent direction.

Hint: Note that $\lambda_{\min}(H)\|u\|\|v\| \leq |u^T H^{-1} v| \leq \lambda_{\max}(H)\|u\|\|v\|$.

Answer: Observe that

$$\tilde{p}^T \nabla\phi = (r - \nabla\phi)^T H^{-1} \nabla\phi$$

and

$$\begin{aligned} r H^{-1} \nabla\phi &\leq \frac{\|r\|\|\nabla\phi\|}{\lambda_{\max}(H)} \\ \nabla\phi^T H^{-1} \nabla\phi &\geq \frac{\|\nabla\phi\|^2}{\lambda_{\min}(H)} \end{aligned}$$

so that

$$\tilde{p}^T \nabla\phi < \frac{\|\nabla\phi\|}{\lambda_{\max}(H)} \left(\kappa(H)\|r\| - \|\nabla\phi\|\right) < 0$$