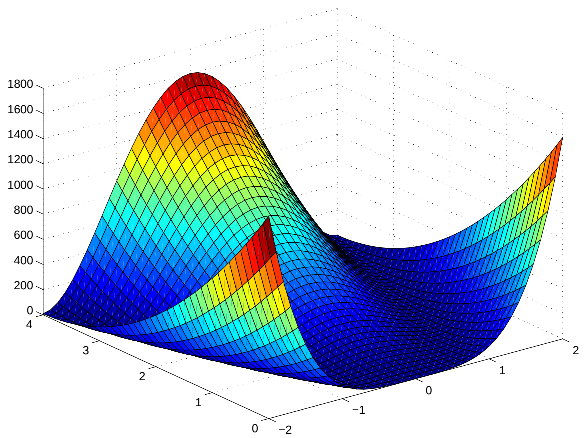
# CS5321 Numerical Optimization Homework 3

Due April 12

1. (50%) The Rosenbrock function $f(x, y) = (1 - x)^2 + 100(y - x^2)^2$ is shown below, whose minimizer is at $(1, 1)$.[1]




   (a) Derive the gradient and the Hessian of $f(x, y)$.

   (b) Read the Matlab code `polyline.m` and `polymod.m` in

   `http://www4.ncsu.edu/~ctk/matlab_darts.html`

   and explain which line search algorithm they are implemented.

   (c) Use $(x_0, y_0) = (-1.2, 1.0)$ to test the steepest descent method with the line search algorithm, implemented in `steep.m` (in the same repository as (b)), and plot its trace $\{(x_k, y_k)\}$. When calling `steep`, the code is like `[x, ...]=steep(x0,@rosenbrock, ...)`. You may modify the code to recode the $\{(x_k, y_k)\}$ and use the plotting code from homework 2.

   (d) Implement Newton's method with line search algorithm, and test it with $(x_0, y_0) = (-1.2, 1.0)$. Plot its trace and compare the results, such as the number of iterations, to (c).

[1] You can find reference of this function in MO and Wikipedia.

2. (25%) We had shown in class that when the line-search algorithm satisfies the Wolfe conditions, $\cos^2\theta_k \|\nabla f_k\|^2 \to 0$, where $\cos\theta_k = \frac{-\vec{p}_k^T \nabla f_k}{\|\nabla f_k\|\|\vec{p}_k\|}$. (Assume $\|\vec{p}_k\| \neq 0$ and $\|\nabla f_k\| \neq 0$.) Therefore, if the search direction of a method satisfies $|\cos\theta_k| > \delta$ for all $k$, then we can prove that $\nabla f_k \to \vec{0}$.

   (a) Assume the matrix norm used satisfies the submultiplicative property, i.e. $\|AB\| \leq \|A\|\|B\|$. Prove that $1/\|\vec{x}\| \geq 1/(\|B\vec{x}\|\|B^{-1}\|)$ for any nonsingular matrix $B$.

   (b) In the Newton-like methods, we replace the Hessian matrix with a symmetric positive definite matrix $B_k$, and use $\vec{p}_k = -B_k^{-1}\nabla f_k$ as the search direction. Use (a) to prove that if $B_k$ is well conditioned, i.e. $\|B_k\|\|B_k^{-1}\| \leq M$ for some constant $M$, then

$$|\cos\theta_k| \geq \frac{1}{M}.$$

   (Hint: you may use the following property directly: For any symmetric positive definite matrix $B$, $\vec{u}^T B\vec{u} \geq 1/\|B^{-1}\|$, where $\vec{u}$ is a unit vector, $\|\vec{u}\| = 1$.)

3. (25%) Prove the formula of SR1.

   (a) Verify the Sherman-Morrison-Woodbury formula.
   For $\hat{A} = A + \vec{a}\vec{b}^T$,

$$\hat{A}^{-1} = A^{-1} - \frac{A^{-1}\vec{a}\vec{b}^T A^{-1}}{1 + \vec{b}^T A^{-1}\vec{a}}.$$

   (Hint: prove $\hat{A}\hat{A}^{-1} = \hat{A}^{-1}\hat{A} = I$.)

   (b) Use (a) and the fact that $B_k$ is symmetric to prove that if $B_k = B_{k-1} + \dfrac{(\vec{y}_k - B_{k-1}\vec{p}_k)(\vec{y}_k - B_{k-1}\vec{p}_k)^T}{(\vec{y}_k - B_{k-1}\vec{p}_k)^T\vec{p}_k}$, then

$$B_k^{-1} = B_{k-1}^{-1} + \frac{(\vec{p}_k - B_{k-1}^{-1}\vec{y}_k)(\vec{p}_k - B_{k-1}^{-1}\vec{y}_k)^T}{\vec{y}_k^T(\vec{p}_k - B_{k-1}^{-1}\vec{y}_k)}.$$