15.084J Recitation Handout 2

Second Week in a Nutshell:

- Newton's Method
- When Newton's Method Fails
- Rates of Convergence
- Quadratic Forms
- Eigenvectors/Eigenvalues/Decompositions

New notation: Q is positive semidefinite is written as $Q \succeq 0$ (and similarly for positive definite, and negative (semi-)definite.

New definition: M is orthonormal if $M^{-1} = M^t$; the rows/columns of M have norm 1 and are perpendicular.

γ is an eigenvalue with eigenvector x of a matrix M if $Mx = \gamma x$

Newton's Method

We can find the solution to a quadratic problem in closed form. Since everything looks quadratic if you squint hard enough, find a quadratic approximation to your problem at a given point, then jump to a new (and presumably better) point which is the minimum of the approximation.

More formally, do a taylor expansion of your function at point x^k... $f(x) \approx f(x^k) + \nabla f(x^k)^t(x - x^k) + \frac{1}{2}(x - x^k)^t H(x^k)(x - x^k)$. Then you can minimize this in closed form to get $x^{k+1} = x^k - H(x^k)^{-1}\nabla f(x^k)$. You can generalize this by doing a line search in the given direction, or moving some multiple of that distance.

When Does Newton's Method Fail

- You can't use the method if you can't get gradients and hessians easily.
- The step is undefined if the Hessian is non-invertable – the function is essentially flat in some direction at the current point
- The method is finding where the gradient is zero – might be a maximum!
- The Hessian must not change too fast (if it does, quadratic approximation is bad)
- The starting point must be "near" an optimum, and you generally don't know how to test for that
- In its current form, it doesn't deal with constraints (though we will later see versions that do)

Rates of Convergence

If we have a sequence x^k converging to $\bar{x}$, we say it has *linear* convergence if $\lim \frac{x^{k+1}-\bar{x}}{x^k-\bar{x}} = \delta < 1$ This means it takes a constant number of iterations to add each significant figure of accuracy.

It has *quadratic* convergence if $\lim \frac{(x^{k+1}-\bar{x})^2}{(x^k-\bar{x})^2} = \delta < 1$ This means it takes a constant number of iterations to double the significant figures of accuracy.

Newton's method has quadratic convergence once it is sufficiently close to a minimum or maximum.

Quadratic Forms

A quadratic form is a function $f(x) = x^tQx + c^tx + d$

Fun facts about quadratic forms:

- You can assume without loss of generality that Q is symmetric
- For minimizing/maximizing, you can assume wolog that d is zero
- The gradient is $Qx + c$
- The Hessian is Q
- The quadratic is convex when $Q \succeq 0$; concave when $Q \preceq 0$

Facts about matricies:

- If Q is real and symmetric, all of its eigenvalues are real, and its eigenvectors are orthogonal
- Thus, you can factor it into $Q = RDR$ where D is the diagonal matrix of eigenvalues, and R is an orthonormal matrix of eigenvectors.
- If $Q \succeq 0$, its eigenvalues are nonnegative; if $Q \succ 0$ its eigenvalues are positive; if $Q \prec 0$ its eigenvalues are negative, and if $Q \preceq 0$ its eigenvalues are nonpositive.
- If Q is symmetric, then $Q \succeq 0$ and nonsingular if and only if $Q \succ 0$.
- If $Q \succeq 0$ then any principal submatrix if Q is $\succeq 0$ (and similarly for $\succ$, $\prec$, and $\preceq$
- If $Q \succ 0$, then $M = \begin{bmatrix} Q & c \\ c^t & b \end{bmatrix}$ has $M \succ 0$ iff $b > c^tQ^{-1}c$