An Example of a Homotopy

This is a supplement to Chapter 24 of *Scientific Computing with Case Studies* by Dianne P. O'Leary, SIAM Press, Philadelphia, 2009

Let's define a homotopy for a convex optimization problem. Suppose we want to find $\boldsymbol{x}_{opt}$ to solve the problem

$$\min_{\boldsymbol{x}} f(\boldsymbol{x})$$

where $f : \mathcal{R}^n \rightarrow \mathcal{R}^1$ is a convex function that is bounded below and has two continuous derivatives. (The boundedness assumption assures that a bounded solution $\boldsymbol{x}_{opt}$ exists.) Also assume that $\boldsymbol{x}_{opt}$ is unique. As in Chapter 9, we denote the gradient of f by $\boldsymbol{g}(\boldsymbol{x})$, and we seek a point for which $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0}$. Now consider the mapping

$$\rho_a(\lambda, \boldsymbol{x}) = \lambda \boldsymbol{g}(\boldsymbol{x}) + (1-\lambda)(\boldsymbol{x} - \boldsymbol{a})\,.$$

Will this homotopy generate a solution path that terminates in $(1, \boldsymbol{x}_{opt})$? We verify some properties of f and the homotopy using the definitions in the previous section.

- Let $U = \mathcal{R}^n \times [0,1) \times \mathcal{R}^n$ and define the function $\boldsymbol{\rho} : U \rightarrow \mathcal{R}^n$ to be $\boldsymbol{\rho}(\boldsymbol{a}, \lambda, \boldsymbol{x}) = \boldsymbol{\rho}_a(\lambda, \boldsymbol{x})$. Let's look at the Jacobian matrix for the mapping. The Jacobian with respect to the $\boldsymbol{x}$ variables is

  $$\boldsymbol{J}_{\boldsymbol{x}} = \lambda \boldsymbol{H}(\boldsymbol{x}) + (1-\lambda)\boldsymbol{I}\,,$$

  where $\boldsymbol{H}$ is the Hessian matrix of $\boldsymbol{f}$. The matrix $\boldsymbol{H}$ is positive semidefinite since f is convex, and adding $(1-\lambda)\boldsymbol{I}$ shifts each of the eigenvalues by $(1-\lambda)$, so $\boldsymbol{J}_{\boldsymbol{x}}$ has rank n for $\lambda \in [0,1)$. Therefore, the complete Jacobian of $\boldsymbol{\rho}$ has full-rank. Therefore, $\boldsymbol{\rho}$ is transversal to $\boldsymbol{0}$ on U.

- The equation $\boldsymbol{\rho}_{\boldsymbol{a}}(0, \boldsymbol{x}) = \boldsymbol{0}$ should have a unique solution $\boldsymbol{x}_0$. Since $\boldsymbol{\rho}_{\boldsymbol{a}}(0, \boldsymbol{x}) = \boldsymbol{x} - \boldsymbol{a}$, this is clearly satisfied, with $\boldsymbol{x}_0 = \boldsymbol{a}$.

Because of these first two properties, we know that the path of zeros of $\boldsymbol{\rho}_{\boldsymbol{a}}$ exists for $\lambda \in [0,1]$. The only bad thing that could happen, then, is that the path wander off to infinity rather than terminating at a finite solution to our original problem. The next few properties show that this cannot happen.

- For this homotopy,

  $$(\boldsymbol{x} - \boldsymbol{x}_{opt})^T \boldsymbol{g}(\boldsymbol{x}) \geq 0. \tag{1}$$

  To see this, recall that by definition, f is a convex function if and only if, for all $\boldsymbol{x}, \boldsymbol{y}$ in the domain of f and all $0 \leq \alpha \leq 1$,

  $$\begin{aligned} f(\alpha \boldsymbol{y} + (1-\alpha)\boldsymbol{x}) &\leq \alpha f(\boldsymbol{y}) + (1-\alpha) f(\boldsymbol{x}) \\ &= \alpha(f(\boldsymbol{y}) - f(\boldsymbol{x})) + f(\boldsymbol{x})\,, \end{aligned}$$

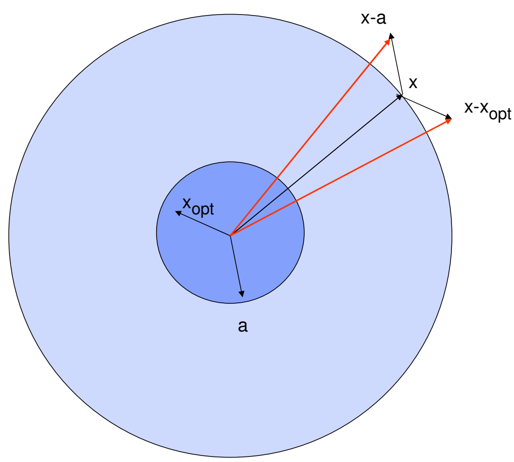

Figure 1. *The geometry of the argument for boundedness of the iterates in the homotopy method for convex optimization. The radius of the outer ball is three times the radius of the inner one. If* $\mathbf{x}_{opt}$ *and* $\mathbf{a}$ *are any points contained in the inner ball, then for any point* $\mathbf{x}$ *on the surface of the outer ball, the inner product between the vectors* $\mathbf{x} - \mathbf{x}_{opt}$ *and* $\mathbf{x} - \mathbf{a}$ *is positive.*

Let $\boldsymbol{z} = \boldsymbol{y} - \boldsymbol{x}$. Then $f(\alpha\boldsymbol{z} + \boldsymbol{x}) = f(\alpha\boldsymbol{y} + (1-\alpha)\boldsymbol{x})$ so the inequality above says that

$$f(\boldsymbol{x} + \alpha\boldsymbol{z}) - f(\boldsymbol{x}) \leq \alpha(f(\boldsymbol{y}) - f(\boldsymbol{x})).$$

Therefore, for $0 < \alpha \leq 1$,

$$\frac{f(\boldsymbol{x} + \alpha\boldsymbol{z}) - f(\boldsymbol{x}) - \alpha\boldsymbol{z}^T\boldsymbol{g}(\boldsymbol{x})}{\alpha} \leq \frac{\alpha(f(\boldsymbol{y}) - f(\boldsymbol{x})) - \alpha\boldsymbol{z}^T\boldsymbol{g}(\boldsymbol{x})}{\alpha}.$$

Taking the limit as $\alpha \to 0$, the left-hand side is zero, so

$$0 \leq (f(\boldsymbol{y}) - f(\boldsymbol{x})) - \boldsymbol{z}^T\boldsymbol{g}(\boldsymbol{x})\,.$$

Set $\boldsymbol{y} = \boldsymbol{x}_{opt}$. Then $f(\boldsymbol{y}) - f(\boldsymbol{x}) \leq 0$, so $-\boldsymbol{z}^T\boldsymbol{g}(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{x}_{opt})^T\boldsymbol{g}(\boldsymbol{x}) \geq 0$ as claimed.

- The solution path is bounded. To show this, choose a number M such that $\|\boldsymbol{x}_{opt}\| < M$ and $\|\boldsymbol{a}\| < M$. Consider the ball defined by $\|\boldsymbol{x}\| = 3M$. For $\boldsymbol{x}$ on this ball, we see from Figure 1 that

$$(\boldsymbol{x} - \boldsymbol{x}_{opt})^T(\boldsymbol{x} - \boldsymbol{a}) > 0,$$

and, because of (1) and the fact that $\boldsymbol{g}(\boldsymbol{x}_{opt}) = \mathbf{0}$, we have

$$(\boldsymbol{x} - \boldsymbol{x}_{opt})^T\boldsymbol{g}(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{x}_{opt})^T(\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{x}_{opt})) \geq 0\,.$$

Put these two inequalities together to get

$$(\boldsymbol{x} - \boldsymbol{x}_{opt})^T \rho_{\boldsymbol{a}}(\lambda, \boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{x}_{opt})^T (\lambda \boldsymbol{g}(\boldsymbol{x}) + (1 - \lambda)(\boldsymbol{x} - \boldsymbol{a})) > 0$$

for $\lambda \in [0, 1)$. So $\boldsymbol{\rho_a}(\lambda, \boldsymbol{x}) \neq \boldsymbol{0}$ on the ball defined by $\|\boldsymbol{x}\| = 3M$. Therefore, the homotopy path, defined by the set of points satisfying $\boldsymbol{\rho_a}(\lambda, \boldsymbol{x}) = \boldsymbol{0}$, cannot intersect the boundary of the $3M$ ball and is therefore forced to remain inside this ball. Therefore the path cannot diverge to infinity.

Putting all of this together, we see that this homotopy generates a solution path that terminates in $(1, \boldsymbol{x}_{opt})$ for almost every $\boldsymbol{a}$.

Note that this example illustrates the theory of homotopy algorithms, and the result means that the homotopy method is one approach to solving convex optimization problems, but such problems are generally too easy to require such heavy machinery. We generally use one of the Newton-type algorithms of Chapter 9 to solve convex optimization problems.