UC: Iterative Solution Algorithms

A. J. Conejo, R. Sioshansi, 2016

THE OHIO STATE UNIVERSITY

What

1. Iterative Solution Algorithms
2. Line Search

What

$$\min_{x \in \mathbb{R}^n} \; f(x),$$

where $f(\cdot) : \mathbb{R}^n \to \mathbb{R}$ is the objective function being minimized.

Iterative Solution Algorithms

Iterative Solution Algorithms

High-level overview of a Generic Algorithm for Unconstrained Nonlinear Optimization Problems.

Iterative Solution Algorithms

```
procedure GENERIC ITERATIVE ALGORITHM
    k ← 0                                          ▷ Set iteration counter to 0
    Fix x^0                                        ▷ Fix a starting point
    while Termination criteria are not met do
        Find direction, d^k, to move in
        Determine step size, α^k
        x^{k+1} ← x^k + α^k d^k                    ▷ Update point
        k ← k + 1                                  ▷ Update iteration counter
    end while
end procedure
```

Iterative Solution Algorithms

In Steps 2 and 3 we initialize the algorithm by setting k, the counter for the number of iterations completed, to 0 and picking a starting point, x^0.

We use the notational convention here that superscripts denote the value of a variable after that number of iterations are completed. Thus, x^0, is a vector of decision variables after 0 iterations have been completed (i.e., it is an initial guess).

Steps 4 through 9 are the iterative procedure.

Iterative Solution Algorithms

In Step 4 we check to see whether we have met certain termination criteria.

If so, we stop and output the incumbent point that we have, x^k.

Otherwise, we proceed with an additional iteration. Each iteration consists of two main procedures.

Iterative Solution Algorithms

In Step 5 we find a **search direction** in which we would like to move away from the incumbent point, x^k.

In Step 6 we conduct what is known as a **line search** to determine how far we would like to move in the search direction found in Step 5. The value, α^k, that we find is called the **step size**

Iterative Solution Algorithms

Finally, in Step 7 we update our point based on the search direction and step size found in Steps 5 and 6 and we update our iteration counter in Step 8.

Iterative Solution Algorithms

1: Find direction, d^k, to move in

2: Determine step size, α^k

3: $x^{k+1} \leftarrow x^k + \alpha^k d^k$

Iterative Solution Algorithms

To use the Generic Algorithm for Unconstrained Nonlinear Optimization Problems, there are thee important details that must be addressed.

The first is how to determine the search direction.

The second is what termination criteria should be used.

The third is how to conduct the line search.

Iterative Solution Algorithms

We start with the last one:

line search.

Line Search

Line Search

Once we have determined a direction to move in (e.g., the minus gradient), the next question is how far to move in the direction we have identified.

Line Search

This is the purpose of a line search. We call this process a line search because what we are doing is looking along a line in the direction that we have identified to move away from the incumbent point.

Looking along this line, we determine how far to move in the direction.

Line Search

An **exact line search** or **line minimization** solves the minimization problem:

$$\min_{\alpha} \; f(x^k + \alpha d^k),$$

to determine the optimal choice of step size.

Note that $f(x^k + \alpha d^k) : \mathbb{R} \to \mathbb{R}$.

Line Search

The advantage of an exact line search is that it provides the absolute best choice of step size.

This comes at a cost, however, which is that we must solve an optimization problem.

This is always a single-variable problem, however, and in some instances the exact line-search problem can be solved relatively easily.

Line Search

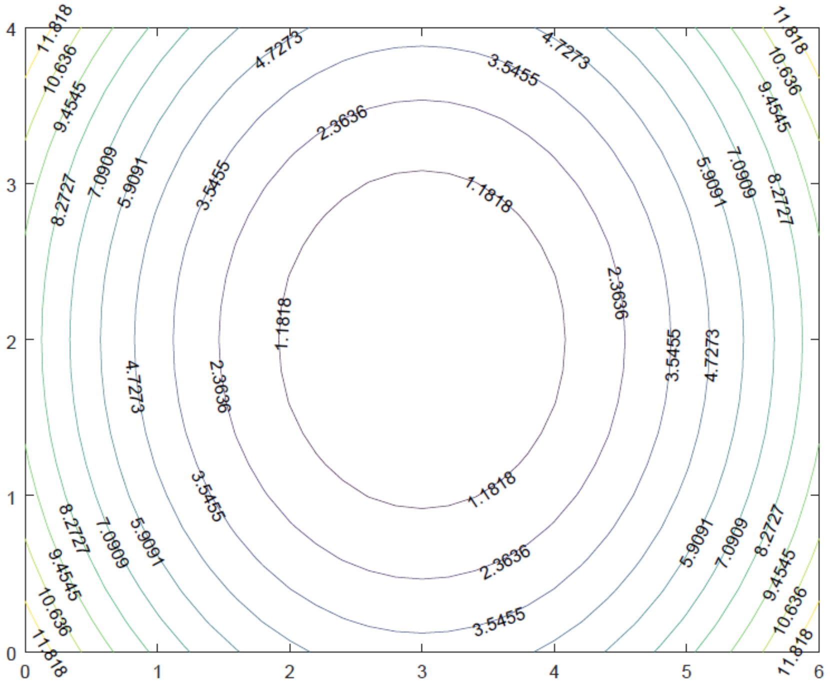
Indeed, the exact line search-problem typically simplifies to solving:

$$\frac{d}{d\alpha} f(x^k + \alpha d^k) = 0,$$

because we can treat the exact line search as an unconstrained optimization problem.

Line Search: Example

Line Search: Example



Consider the unconstrained problem:

$$\min_{x} \ f(x) = (x_1 - 3)^2 + (x_2 - 2)^2.$$

We start from the point $x^0 = (1, 1)^\top$ and the search direction:

$$d^0 = -\nabla f(x^0) = \begin{pmatrix} 4 \\ 2 \end{pmatrix}.$$

Line Search: Example

To conduct an exact line search we solve the following minimization problem:

$$
\begin{aligned}
\min_{\alpha^0} \; f\left(x^0 + \alpha^0 d^0\right) &= f\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha^0 \begin{pmatrix} 4 \\ 2 \end{pmatrix}\right) \\
&= f\left(\begin{pmatrix} 1 + 4\alpha^0 \\ 1 + 2\alpha^0 \end{pmatrix}\right) \\
&= (4\alpha^0 - 2)^2 + (2\alpha^0 - 1)^2.
\end{aligned}
$$

Line Search: Example

To solve this unconstrained minimization, we use the FONC which is:

$$\frac{d}{d\alpha^0}\left[(4\alpha^0 - 2)^2 + (2\alpha^0 - 1)^2\right] = 8(4\alpha^0 - 2) + 4(2\alpha^0 - 1) = 0,$$

which gives $\alpha^0 = 1/2$.

Line Search: Example

We further have that:

$$\frac{d^2}{d{\alpha^0}^2}\left[(4\alpha^0-2)^2+(2\alpha^0-1)^2\right]=40>0,$$

meaning that this value of α^0 is a global minimum. Thus, our new point is:

$$x^1=x^0+\alpha^0 d^0=(3,2)^\top.$$

Line Search: Alternatives to the exact search

1. Quadratic fit.
2. Cubic fit.

Line Search: Quadratic fit

Line Search: Quadratic fit

Consider 3 points: $(\alpha_1, f_1), (\alpha_2, f_2), (\alpha_3, f_3)$

The minimizer of the quadratic function passing throuhg these points is:

$$\alpha^* = \frac{1}{2}\frac{b_{23}f_1 + b_{31}f_2 + b_{12}f_3}{a_{23}f_1 + a_{31}f_2 + a_{12}f_3}$$

where: $a_{ij} = \alpha_i - \alpha_j$ and $b_{ij} = \alpha_i^2 - \alpha_j^2$, $i, j = 1, 2, 3, i \neq j$

Upate the current point as: $x^{k+1} \leftarrow x^k + \alpha^* d^k$

Line Search: Cubic fit

Line Search: Cubic fit

Consider 2 points and its derivatives: $(\alpha_1, f_1), (\alpha_2, f_2), f_1', f_2'$

The minimizer of the cubic function fitting these points is

$$\alpha^* = \alpha_2 - (\alpha_2 - \alpha_1)\left[\frac{f_2' + u_2 - u_1}{f_2' - f_1' + 2u_2}\right]$$

where:

$$u_1 = f_1' + f_2' - 3\frac{f_1 - f_2}{\alpha_1 - \alpha_2} \text{ and } u_2 = \sqrt{u_1^2 - f_1' f_2'}$$

Upate the current point as: $x^{k+1} \leftarrow x^k + \alpha^* d^k$

This is it!