Theory for Boundary-Value Objective-First Optimization

Jesse Lu, jesselu@stanford.edu

June 13, 2011

Contents

1 Improving the field

We begin by writing down a generic (sourceless) physics problem,

$$A(p)x = 0. \tag{1}$$

Here, x is the field variable, p is the structure variable, and $A(p)$ represents the physics of the problem.

Now, suppose that we want to fix the value of the field at certain grid-points. We denote these fixed points as *boundary values* of the problem. Specifically, we break up x into

$$x = S_1 x_1 + S_0 x_0, \tag{2}$$

where S_1 and S_0 are selection matrices for the varying and fixed elements of x respectively.

We can now attempt to solve eq. 1 by finding x_1 in the following manner. Let

$$A(p)S_1 x_1 = -S_0 x_0, \tag{3}$$

or

$$\hat{A}(p)x_1 = b, \tag{4}$$

where $\hat{A}(p) = A(p)S_1$ and $b = -S_0 x_0$.

Note that there often will not be a valid x_1 to satisfy eq. [**?**] since $\hat{A}(p)$ will in general be skinny and full-rank. Instead, we can minimize the *physics residual*, defined as

$$\|A(p)x\|^2 = \|\hat{A}(p)x_1 - b\|^2. \tag{5}$$

We call solving the problem

$$\underset{x_1}{\text{minimize}} \quad \|\hat{A}(p)x_1 - b\|^2 \tag{6}$$

improving the field. Note that an equivalent problem definition is

$$\begin{aligned} &\underset{x}{\text{minimize}} && \|A(p)x\|^2 && (7)\\ &\text{subject to} && S_0^T x = x_0. && (8) \end{aligned}$$

2 Improving the structure

We now consider the structure improvement problem, that is,

$$\underset{p}{\text{minimize}} \quad \|A(p)x\|^2. \tag{9}$$

Whereas the field improvement problem can always be solved exactly, this is not the case for the structure improvement problem. Even when $A(p)$ is linear with respect to p, there are often restrictions on p which make finding a solution very difficult. For example, a common restriction is that p be binary, that is, $p \in p_0, p_1$. In this case, the problem is generally NP-hard. For these reasons, we often use a simple gradient-descent method to arrive at an approximate solution of eq. 9.

If $A(p) = A_1 \text{diag}(A_2 p) A_3$ then,

$$\begin{aligned} A(p)x &= A_1 \text{diag}(A_2 p) A_3 x && (10)\\ &= A_1 \text{diag}(A_3 x) A_2 p && (11)\\ &= B(x)p. && (12) \end{aligned}$$

where $B(x) = A_1 \text{diag}(A_3 x) A_2$.

The gradient of the physics residual with respect to p can then be computed using

$$\frac{\partial}{\partial p} \frac{1}{2} \|A(p)x\|^2 = \frac{\partial}{\partial p} \frac{1}{2} \|B(x)p\|^2 = B(x)^* B(x)p. \tag{13}$$

3 Gradients

A first-order approximation to a function, $f(x) : \mathcal{C}^n \to \mathcal{R}$ can be formulated as

$$f(x) \approx f(x_0) + \mathcal{R}\{g(x_0)^*(x - x_0)\}, \tag{14}$$

where $g(x_0)$ is the gradient of f at x_0.

To test that $g(x)$ does indeed give the correct value for the gradient, one can produce a set random of vectors v_n and find the error,

$$\frac{\|(f(x_0 + v_n) - f(x_0)) - \mathcal{R}\{g(x_0)^*(v_n)\}\|^2}{\|v_n\|^2} \tag{15}$$

for $\|v_n\|^2 \ll 1$.

References