1 DPG Method for Non Linear Problems

We are interested in solving a non-linear variational problem,

$$\left\{ \begin{array}{l} u \in U \\ b(u;v) = l(v) \quad v \in V \end{array} \right. \tag{1.1}$$

where U, V are two Hilbert spaces, form $b(u;v)$ is (anti)linear in v but non-linear in u. Introducing the non-linear operator $B : U \to V'$ corresponding to $b(u,v)$,

$$\langle B(u), v \rangle_{V' \times V} = b(u;v) \quad u \in U, v \in V \tag{1.2}$$

we can rewrite the problem as

$$B(u) = l \quad \text{ in } V' \tag{1.3}$$

Let $U_h \subset U$ be a finite-dimensional subspace. The minimum residual method seeks a solution that minimizes the residual,

$$J(u_h) := \frac{1}{2} \| R_V^{-1} (B(u_h) - l) \|_V^2 \to \min_{u_h \in U_h} \tag{1.4}$$

where $R_V : V \to V'$ is the Riesz operator.

The purpose of this note is to discuss the computation of the first and the second Gatéaux derivatives of function $J(u_h)$.

First derivative. We have

$$\begin{aligned} \langle J'(u_h); \delta u_h \rangle \quad &= (R_V^{-1}(B(u_h) - l), R_V^{-1} B'(u_h; \delta u_h)_V \\ &= \langle B(u_h) - l, R_V^{-1} B'(u_h; \delta u_h) \rangle \\ &= \langle r_h, v_{\delta u_h} \rangle \end{aligned} \tag{1.5}$$

where $B'(u;\delta u)$ denotes the Gateaux derivative of operator B at u, linear in δu, $r_h := B(u_h) - l \in V'$ is the residual, and $v_{\delta u_h}$ is the test function cprresponding to variation δu_h obtained by inverting the Riesz operator,

$$\left\{ \begin{array}{l} v_{\delta u_h} \in V \\ (v_{\delta u_h}, \delta v)_V = \langle B'(u_h; \delta u_h), \delta v \rangle \quad \delta v \in V \end{array} \right. \tag{1.6}$$

Equivalently, we can use the notation with scalar-valued forms,

$$\langle B'(u_h; \delta u_h), \delta v \rangle = b'(u_h; \delta u_h; \delta v) \quad u_h, \delta u_h \in U_h, \delta v \in V \tag{1.7}$$

Second derivative. Differentiating (1.5), we get

$$\langle J''(u_h);\delta u_h,\Delta u_h\rangle = (R_V^{-1}B'(u_h;\Delta u_h),R_V^{-1}B'(u_h;\delta u_h))_V + (R_V^{-1}(B(u_h)-l),R_V^{-1}B''(u_h;\delta u_h,\Delta u_h))_V \tag{1.8}$$

where

$$B''(u_h;\delta u_h,\Delta u_h) = b''(u_h;\delta u_h,\Delta u_h;\cdot) \in V' \tag{1.9}$$

is the second derivative of operator B evaluated at u_h in the directions δu_h and Δu_h. The first term can be rewritten as

$$(R_V^{-1}B'(u_h;\Delta u_h),R_V^{-1}B'(u_h;\delta u_h))_V = (v_{\Delta u_h},v_{\delta u_h})_V = b'(u_h;\Delta u_h;v_{\delta u_h}) \tag{1.10}$$

which indicates the way it should be computed for the DPG method in practice. If N is dimension of the *enriched space* for an element, inversion of the Gram matrix in (1.6) using Cholesky decomposition is worth $O(N^3)$ operations but the computation of the right-hand side in (1.10) is worth only $O(N^2)$ operations and, therefore, its cost is neglible.

Introducing the Riesz representation of the residual,

$$\delta v_h := R_V^{-1}\overline{r_h} = R_V^{-1}(\overline{B(u_h)-l}) \tag{1.11}$$

we can rewrite the second terms as

$$(R_V^{-1}r_h,R_V^{-1}B''(u_h;\delta u_h,\Delta u_h))_V = (R_V^{-1}B''(u_h;\delta u_h,\Delta u_h),R_V^{-1}\overline{r_h})_V = \langle B''(u_h;\delta u_h,\Delta u_h),\delta v_h\rangle \tag{1.12}$$

Notice that computation of (1.11) is worth one more resolution of (1.6) with the right-hand side replaced by the (conjugated) residual.

The final formula for the second derivatives is as follows,

$$\langle J''(u_h);\delta u_h,\Delta u_h\rangle = b'(u_h;\Delta u_h;v_{\delta u_h}) + b''(u_h;\delta u_h,\Delta u_h;\delta v_h) \tag{1.13}$$

Upgrades to the linear code.

1. Computation of first and second derivatives of the scalar-valued form $b(u;v)$:

   $$b'(u;\delta u;v), \quad b''(u;\delta u,\Delta u;v) \tag{1.14}$$

   This should be done in two new, separate routines.

2. Modification of the computation of the optimal test functions involving:

   (a) replacement of $b(\delta u_h,\delta v)$ in the right-hand side of the local problem with $b'(u_h;\delta u_h;\delta v)$;

   (b) computation of element residual r_h;

   (c) one extra resolution step to compute δv_h.

3. Computation of element contributions to gradient and hessian using formulas (1.5) and (1.13).