General notes on DPG

March 3, 2014

1 Least squares

We begin with a general variational formulation

$$b(u,v) = l(v)$$

DPG begins with the idea that you would like to do least squares on the operator equation

$$Bu = \ell, \quad Bu, \ell \in V'$$

where $\langle Bu, v\rangle_{V'\times V} = b(u,v)$ and $\langle \ell, v\rangle_{V'\times V} = l(v)$. Since $Bu - \ell \in V'$, we minimize the norm of this residual in V' over the finite dimensional space U_h, i.e.

$$\min_{u_h \in U_h} \|Bu_h - \ell\|_{V'}^2 .$$

This leads to the normal equations

$$\left(Bu - \ell, B\delta u\right)_{V'}, \quad \forall \delta u \in U_h.$$

The Riesz map gives us the equivalent definition

$$\left(R_V^{-1}\left(Bu - \ell\right), R_V^{-1}\left(B\delta u\right)\right)_V = 0, \quad \forall \delta u \in U_h.$$

Assuming we've specified the Riesz map through a test space inner product

$$\langle R_V v, \delta v\rangle_{V'\times V} = (v, \delta v)_V,$$

this leads to what I call a Dual Petrov-Galerkin method.

2 Algebraic perspective

In the above example, V is infinite dimensional. If we approximate V by V_h such that $\dim(V_h) > \dim(U_h)$, we get matrix representations of our operators

$$\begin{aligned} B_{ij} &= b(u_j, v_i), \quad u_j \in U_h, v_i \in V_h \\ R_V &= (v_i, v_j)_V \quad v_i, v_j \in V_h \\ \ell_i &= l(v_i), \quad v_i \in V_h. \end{aligned}$$

The resulting normal equations

$$\left(R_V^{-1}\left(Bu - \ell\right), R_V^{-1}\left(B\delta u\right)\right)_V, \quad \forall \delta u \in U_h.$$

can now be written as

$$\left(R_V^{-1}\left(Bu-\ell\right)\right)^T R_V \left(R_V^{-1}B\right) = 0,$$

or, after simplifying to $\left(Bu-\ell\right)^T R_V^{-1}B = 0$, we get the algebraic normal equations

$$B^T R_V^{-1} B u = B^T R_V^{-1} \ell.$$

This is just the solution to the algebraic least squares problem

$$\min_u \|Bu-\ell\|^2_{R_V^{-1}}.$$

Such problems can also be written using the augmented system for the least squares problem

$$\begin{bmatrix} R_V & B \\ B^T & 0 \end{bmatrix} \begin{bmatrix} e \\ u \end{bmatrix} = \begin{bmatrix} l \\ 0 \end{bmatrix}.$$

This can be interpreted as the mixed form of the Dual Petrov-Galerkin method, which is used by Cohen, Welper, and Dahmen in their 2012 paper "Adaptivity and variational stabilization for convection-diffusion equations".

$$\begin{aligned} (e,v)_V + b(u,v) &= l(v) \\ b(\delta u, e) &= 0 \end{aligned}$$

Eliminating e from the above system leads to the above algebraic normal equations.

Note: for general R_V, the algebraic normal equations are completely dense. Cohen, Welper, and Dahmen thus solve the augmented system to get solutions in this setting; however, this is a saddle point problem, and over 2x as large as the trial space, which makes preconditioning and solving more difficult.

3 Deriving the Discontinuous Petrov-Galerkin method

We want to avoid solving either a fully dense system or a saddle point problem, so we introduce Lagrange multipliers $\widehat{u}$ to enforce continuity weakly on e, which we will now approximate using discontinuous functions. These $\widehat{u}$ are defined on element edges only, similarly to hybrid variables or mortars in finite elements. This leads to the new system

$$\begin{aligned} \langle \widehat{u}, [\![v]\!] \rangle_{\Gamma_h} + (e,v)_V + b(u,v) &= l(v) \\ b(\delta u, e) &= 0 \\ \langle \hat{\mu}, [\![e]\!] \rangle_{\Gamma_h} &= 0 \end{aligned}$$

where Γ_h is the mesh skeleton (union of all element edges). The resulting algebraic system here is

$$\begin{bmatrix} R_V & B & \hat{B} \\ B^T & 0 & 0 \\ \hat{B}^T & 0 & 0 \end{bmatrix} \begin{bmatrix} e \\ u \\ \widehat{u} \end{bmatrix} = \begin{bmatrix} f \\ 0 \\ 0 \end{bmatrix}.$$

Because e is now discontinuous, R_V is made block-diagonal; eliminating e returns the (fairly) sparse symmetric positive-definite DPG system.