General notes on DPG

July 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

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

The Riesz map gives us the equivalent definition

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

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

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

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 $(Bu - \ell)^T R_V^{-1} B = 0$, we get the algebraic normal equations

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

This is just the solution to the algebraic least squares problem

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

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".

$$(e, v)_V + b(u, v) = l(v)$$
$$b(\delta u, e) = 0$$

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

$$\langle \widehat{u}, \llbracket v \rrbracket \rangle_{\Gamma_h} + (e, v)_V + b(u, v) = l(v)$$
$$b(\delta u, e) = 0$$
$$\langle \hat{\mu}, \llbracket e \rrbracket \rangle_{\Gamma_h} = 0$$

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

$$A = \begin{bmatrix} B^T R_V^{-1} B & B^T R_V^{-1} \hat{B} \\ \hat{B}^T R_V^{-1} B & \hat{B}^T R_V^{-1} \hat{B} \end{bmatrix} \begin{bmatrix} u \\ \widehat{u} \end{bmatrix} = \begin{bmatrix} B^T R_V^{-1} f \\ \hat{B}^T R_V^{-1} f \end{bmatrix}.$$

3.1 Primal hybrid formulation: why it works for DPG

Note: primal hybrid formulation gives Crouziex-Raviart elements for lowest order $N = 1$, with flux order $N_f = N - 1$. However, this does not work in defining elements for $N > 1$; for N even and $N_f = N - 1$, the Lagrange multiplier constraint set will not be linearly independent. On triangles, the jump constraints for N odd will imply that the solution must be an orthogonal polynomial of order N over the edge. Since N is odd, in 2D, the values at both vertices will be equal, and the trace on the edge will be continuous. Taking u to be any polynomial over T with this trace gives that $Bu = 0$, but $u \neq 0$. (I believe Jay Gopalakrishnan also noticed this independently.)

If you take N arbitrary and $N_f \leq N - 2$, however, things work fine. This is roughly the situation with DPG - consider primal DPG, with N_{trial}, $N_{\text{test}} = N_{\text{trial}} + 2$, and $N_f = N_{\text{trial}} - 1$.

4 Primal DPG and preconditioning

The DPG system can be reordered

$$A = \begin{bmatrix} B^T R_V^{-1} B & B^T R_V^{-1} \hat{B} \\ \hat{B}^T R_V^{-1} B & \hat{B}^T R_V^{-1} \hat{B} \end{bmatrix} = \begin{bmatrix} K & C \\ C^T & F \end{bmatrix}$$

where K is the matrix corresponding to field (volume) variables, F is the submatrix corresponding to flux (surface/mortar) variables, and C couples the system together.

We can do a block triangular factorization of this system

$$A = \begin{bmatrix} I & 0 \\ C^T K^{-1} & I \end{bmatrix} \begin{bmatrix} K & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} I & K^{-1}C \\ 0 & I \end{bmatrix}.$$

where $S = F - C^T K^{-1} C$, the Schur complement. We approach the design of preconditioners based on the approximate inversion of K and S.

For a first pass, we consider the primal DPG method, which is probably the simplest DPG method available, where field variables are approximated with C_0-continuous piecewise polynomials and the flux variables are approximated using discontinuous mortar basis functions. Test functions are approximated with disjoint discontinuous polynomials of higher order than the trial field variables.

We may also wish to consider the DPG method with ultra-weak variational formulation (discontinuous field variables). In this case, K is also block diagonal.

4.1 K block

We assume that preconditioning the K submatrix can roughly be done using similar techniques to preconditioning standard elliptic equations. Initial numerical evidence appears to support this

- Two-level additive Schwarz. A coarse solve (P_1 elements with AGMG?) + (overlapping) block solves. Not working yet.

- Direct AGMG preconditioning. Working for CG and DPG.

- P_1 FEM preconditioning of nodal bases Working for CG, not yet implemented for DPG.

4.2 Schur block

Preconditioners for this general sort of system often involve an approximation of the Schur complement, which, in this case, is

$$S = \hat{B}^T R_V^{-1} \hat{B} - \hat{B}^T R_V^{-1} B \left(B^T R_V^{-1} B \right)^{-1} B^T R_V^{-1} \hat{B}$$

Factoring out $\hat{B}$ on both sides gives

$$S = \hat{B}^T \left(R_V^{-1} - R_V^{-1} B \left(B^T R_V^{-1} B \right)^{-1} B^T R_V^{-1} \right) \hat{B}$$

Factoring out $\hat{B}^T R_V^{-1}$ on both the left and right gives

$$S = \hat{B}^T R_V^{-1} \left(R_V - B \left(B^T R_V^{-1} B \right)^{-1} B^T \right) R_V^{-1} \hat{B}$$

If we define $\tilde{B} = R_V^{-1/2} B$, then the above is

$$S = \hat{B}^T R_V^{-1/2} \left[I - \tilde{B}(\tilde{B}^T \tilde{B})^{-1} \tilde{B}^T \right] R_V^{-1/2} \hat{B}.$$

The matrix $\left[I - \tilde{B}(\tilde{B}^T \tilde{B})^{-1} \tilde{B}^T \right]$ is the orthogonal projector onto the kernel of $\tilde{B}^T$ - does this help?

Alternatively, suppose we just wish to precondition S with $C = \hat{B}^T R_V^{-1} \hat{B}$. This assumes that $C^{-1}S = I - C^{-1} B^T A^{-1} B$ is close to an identity, or that $C^{-1} B^T A^{-1} B$ is nearly zero. Numerical experiments with Poisson/convection-diffusion indicate this may holds independently of h, though convection-diffusion appears to have dependence on N for $N > 4$.

4.3 Block Jacobi

We can write (overloading notation for A) the full DPG system in block form

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} u \\ \hat{u} \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}.$$

We could naively break the system up into two blocks and use a fixed point iterative scheme

$$u^{n+1} = A^{-1}(f - B\hat{u}^n)$$
$$\hat{u}^{n+1} = C^{-1}(g - B^T u^{n+1}).$$

or equivalently

$$\begin{bmatrix} u \\ \hat{u} \end{bmatrix} = \begin{bmatrix} A^{-1} & \\ & C^{-1} \end{bmatrix} \left(\begin{bmatrix} f \\ g \end{bmatrix} - \begin{bmatrix} & B \\ B^T & \end{bmatrix} \begin{bmatrix} u \\ \hat{u} \end{bmatrix} \right)$$

which we can recognize as a block Jacobi iteration. This may be preferrable in a matrix-free environment, as both A and C can be computed in a matrix-free fashion ($C\hat{u}$ can be computed using local inversions of the Riesz product with $\hat{u}$ as boundary data). As far as I can tell, the Schur complement $S = C - B^T A^{-1} B$ for a general DPG system cannot.

This iteration was tested on pure convection under both the ultra-weak and primal formulations. As there was no observable difference in the primal formulation, we just show ultra-weak results here. Test norm was taken to be $\|v\|_V^2 = \alpha \|v\|_{L^2(\Omega)}^2 + \|\beta \cdot \nabla v\|_{L^2(\Omega)}^2$.

- Convergence depends on ΔN and N_{flux}.

- If $N_{\text{flux}} = 0$, convergence is roughly independent of ΔN. If $N_{\text{flux}} > 0$, the convergence is both slower than with $N_{\text{flux}} = 0$ and more sensitive to N and ΔN.

Convergence of the fixed point solution to the exact solution is shown here. We note that these results are sensitive also to the regularization parameter α — smaller α results in faster convergence, but obviously worse conditioning of the local problem. Idea: filtered mass matrix?

Questions:

- Do the results carry over for the ultra-weak formulation? Maybe only using the graph norm?

(a) $N_{\text{flux}} = 1, \Delta N = 2$ (b) $N_{\text{flux}} = 1, \Delta N = 4$

(c) $N_{\text{flux}} = 0, \Delta N = 4$ (d) $N_{\text{flux}} = 0, \Delta N = 2$
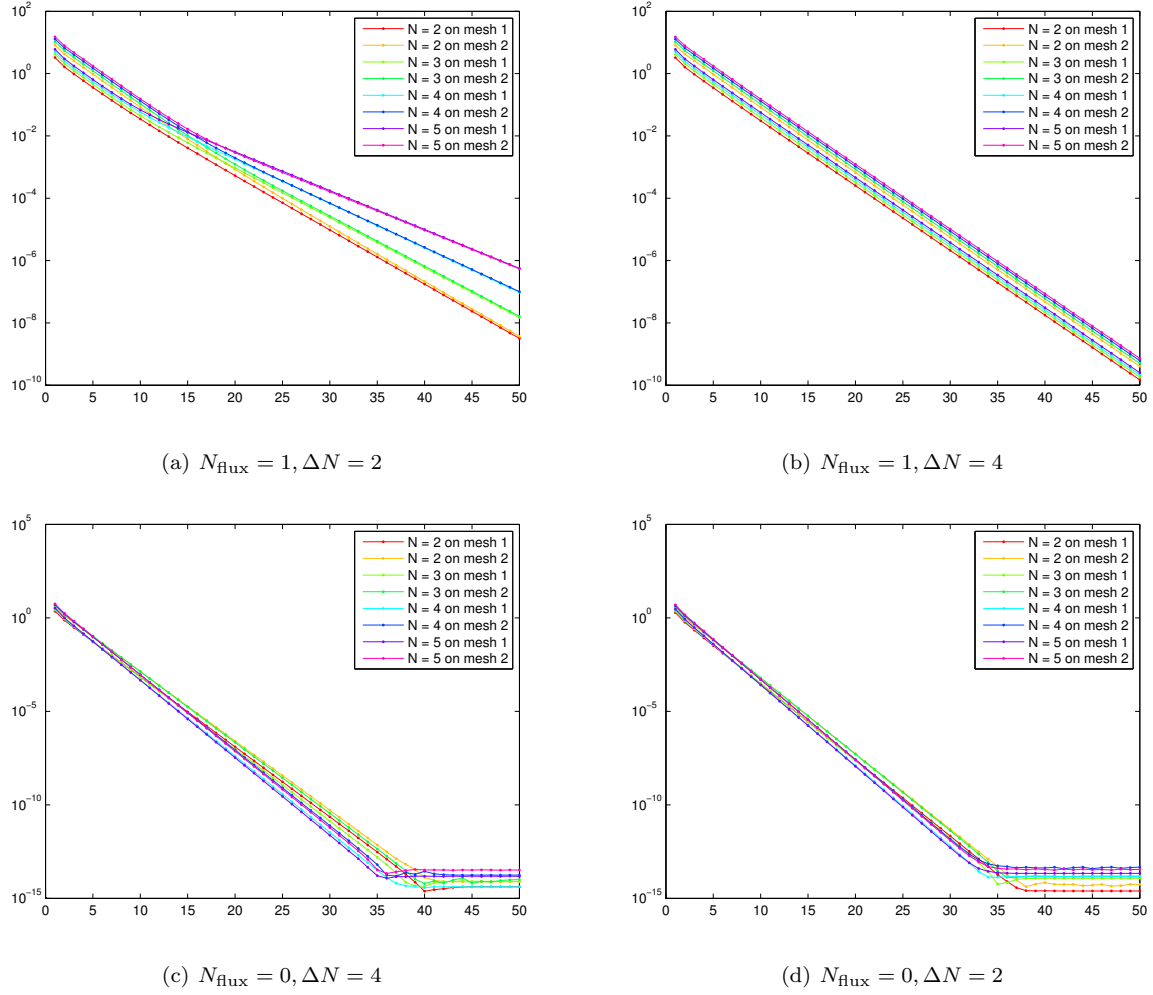
Figure 1: Block Jacobi fixed point iterations for pure convection with $\beta = (1, 0)$, ultra-weak formulation.

- Why do these results fail for Poisson's equation in the primal formulation?

- **This has to be related to the fixed point/Uzawa iteration of Dahmen.** Issues in generalization:
  - Addition of regularizing term α to graph norm (can we duplicate Dahmen's proof under α?)
  - Addition of fluxes (nonconformity).

4.3.1 Relation to Dahmen fixed point iteration

4.4 Etc notes

- Flux sub-block
$$\hat{B} R_V^{-1} \hat{B}$$
  is the exact same as system for primal hybrid method under the test norm. Block preconditioner?

- Can we use the idea of reducing to field dofs? i.e. turn the inversion of the flux matrix into just an outer product? Make equiv to testing with null space of $\hat{B}^T$...

5

- Can we change the coupling to the mortar space for e? i.e. add consistency/penalty/etc?

- Mesh dependent test norms: do they improve the flux subproblem?

- Calderon preconditioners?