FOSLoSophy

Informal discussion of some advantages and disadvantages of First-Order System Least Squares (FOSLS).

As with any new development that breaks from convention, the least-squares approach can be easily misunderstood. Both understatement and overstatement are natural tendencies. Newcomers are quick to recognize the subtlety of this methodology, which is true enough, but it is easy to misjudge its applicability or complexity. It seems at first glance to require increased regularity or smoothness, and to demand more computational resources, but these concerns are usually dispelled on closer inspection. On the other hand, FOSLS is not about to completely replace any other discretization methodology. It is simply a new tool in the struggle for effective solution of partial differential equations. It is compelling, however, in its potential to yield uniformly fast and accurate solvers for some classes of applications involving high Reynolds flow, linear elasticity approaching the incompressible limit, and highly indefinite Helmholtz problems. In any event, the purpose of this note is to start a dialogue that will hopefully lead to a better understanding of where the least-squares approach fits into the field of computational mathematics. We assume that the reader is familiar with basic FOSLS concepts (see the end of this document for relevant sources), but not necessarily with its implications or basic philosophy. We begin with a general discussion and end with an example that is simplified for clarity, but hopefully significant enough to illustrate various points of this discussion.

One of the main advantages of FOSLS is its formulation of a *minimization principle* for problems that have no other natural optimization form. This leads to several benefits: discretization can be done by Rayleigh-Ritz techniques, which are usually simpler than Galerkin methods because they need only one type of subspace and more powerful because they can exploit the optimization structure; analogously, the only basic choices needed for designing multigrid solvers for FOSLS discretizations are relaxation and interpolation (other formulations usually also require choices for restriction and coarse grid operators); such 'variational' multigrid methods are typically more effective and better supported by theory than non-variational schemes; heuristics and analysis of discretization can be based on approximation theory instead of more cumbersome and often misleading approaches like truncation error analysis; and the problem is imbued with a sense of optimality, which is important in developing robust discretizations, especially in the presence of adaptive refinement, singularities and discontinuities, nonlinearities, and widely varying coefficient. Another advantage, which reflects the primary goal of this methodology, is that FOSLS represents a top-down *systematic* approach, with the basic aim of formulating the original problem so well that the numerical process (discretization and multigrid solution) becomes as straightforward and optimal as possible. In this vein, an attempt is made to formulate the functional so that the system variables (e.g., potentials and fluxes) are essentially *decoupled,* meaning that the homogeneous form is *elliptic and continuous* (more generally, it is equivalent to a diagonal form of scalar norms). This allows essentially separate treatment (more precisely, separate consideration) of each variable in the discretization and multigrid solution processes. For example, it avoids restrictions like the LBB condition and staggered grids. When the functional is designed to be fully vector H^1 equivalent, then standard discretization and multigrid solution techniques can be easily and effectively applied. For example, no special 'stabilization' techniques like artificial diffusion or upwinding are needed. Another important attribute is that the FOSLS functional provides a practical and sharp *a posteriori error measure* at no additional cost: the value of the functional provides a measure of both the absolute or the relative total error for any approximation, no matter how it was obtained.

Of course, these advantages must come at some cost. First, there are *more system variables.* For example, a diffusion-convection equation would require four dependent variables in three-dimensional geometry. On the other hand, this does not mean more unknowns or more computational cost, especially if FOSLS strengthened accuracy properties lead to smaller or, generally, more easily solved algebraic systems. Second, to obtain full vector H^1 equivalence (equivalence of the homogeneous form to a form of decoupled H^1 scalar norms), the problem must exhibit *more regularity* and smoothness than it may possess. However, the H^{-1} norm versions of FOSLS apply whenever, say, mixed methods do, and the L^2 norm versions generally require only the additional assumption that the source be in L^2; moreover, these difficulties are usually isolated phenomena, occurring near rough boundaries or internal interfaces that require special treatment in any case. Third, FOSLS formulations are generally *non-conservative,* although special treatment can lead to fully conservative forms. Finally, development of a fully vector-H^1-equivalent FOSLS functional can be *tricky:* augmenting the system with correct equations and boundary conditions can require much understanding and a bit of luck.

FOSLS H^{-1} Norm Version

The H^{-1} norm versions of FOSLS usually involve the simplest reduction of the original equations to first-order form and use of an inverse type norm that in essence preconditions the 'divergence' equations (those equations that usually come from higher derivatives in the original equations) and allows for less smoothness (e.g., of the source terms). As such, this version is usually *simpler* to define and more *generally applicable.* In fact, it applies under the same general framework as do mixed methods. The aim of H^{-1} norm versions of FOSLS is to develop a functional whose homogeneous part is equivalent to a diagonal form, usually involving lower-order norms on the less smooth variables (e.g., fluxes). An issue that has apparently not yet been seriously exploited is the possibility of using norms other than those defined by the adjoint of the H^1 inner product, possibly with variable coefficients for better 'preconditioning' and higher powers of the inverse for broader applicability.

One difficulty with this version of FOSLS stems from the use of these weaker norms: while they do allow for less smooth data, they also generally require *delicate* *balance* between the different terms of the functional. This can be seen by realizing that the primary objective in establishing diagonal equivalence is to show that the off-diagonal terms are small in some sense: for H^{-1} norm versions of FOSLS, these terms are seldom of lower order than the diagonal terms, so there is little leeway to scale the terms of the functional differently. For example, applied to the usual elliptic equation with a diffusion coefficient that varies possibly with jumps and/or large changes of scale over the region, straightforward application of this version of FOSLS does not give diagonal equivalence uniformly in coefficient variations, the symptoms of which are degrading approximation and solver performance. Another aspect of this sensitivity to scaling is the unfeasibility of 'two-stage' schemes. For Poisson's equation using the L^2 version of FOSLS with curl conditions, for example, the flux equations can be weighted with arbitrarily small positive parameters, without loss of accuracy or solver speed, because this FOSLS system is almost lower triangular. In fact, the weights can be set to zero and a proper functional for the fluxes remains, allowing for a second-stage recovery of the potential. Unfortunately, the intimate coupling of the divergence and flux terms prevents this in the H^{-1} norm version of FOSLS. By itself this is not necessarily troublesome, but this sensitivity to scaling might be limiting in the design of parameter independent formulations, for example, in high Reynolds number fluid flow problems, where flexibility of scaling is the basic tool used to devise L^2-type FOSLS functionals that are uniformly effective over all Reynolds number ranges.

FOSLS L^2 Norm Version

When it applies, the L^2 norm version is usually the most effective form of FOSLS. To be sure, H^2 regularity is usually needed for full vector H^1 equivalence of special L^2 norm formulations, and this can be tricky to achieve. However, no additional regularity is really needed to apply the L^2 norm version: it is generally enough that the source be in L^2. The basic advantage of this version is the *stronger sense of well-posedness:* applied to scalar second-order equations, for example, the homogeneous form of the L^2 norm functional exhibits $(H^{div} \bigcap H^{curl}) \times H^1$ equivalence in general, which is locally vector H^1, and it exhibits full vector H^1 equivalence for special forms when the original problem has increased regularity. This equivalence leads to several other benefits, including *optimal performance* of standard discretization and multigrid schemes, some *scale insensitivity* that allows 'two-stage' methods and design of uniformly well-posed functionals for many parameter-dependent problems, and *lower-order variable coupling* that allows some leeway in the approximation of individual variables. We will attempt to clarify some of these benefits in the next section. Finally, as with most formulations of this type that admit efficient multigrid solvers, this approach is highly vectorizable and parallelizable, especially since the individual variables are essentially decoupled and can therefore be processed simultaneously.

Illustration

For concreteness, consider the convection-diffusion equation

$$\nabla^* a \nabla p + \mathbf{b} \cdot \nabla p = f$$

with appropriate boundary conditions, which we ignore here in deference to a simpler, more formal discussion. Here, $\nabla^* = - \nabla \cdot$, a is a scalar function (which without loss of generality we assume satisfies $a \geq 1$), and $\mathbf{b}$ is a vector function. A natural H^{-1} norm FOSLS functional is

$$F\ (\mathbf{u}, p; f) \equiv || \nabla^* \mathbf{u} + \mathbf{b} \cdot (\mathbf{u}/a) - f||_{-1}^2 + ||\mathbf{u}/\sqrt{a} - \sqrt{a} \nabla p||^2,$$

where $|| \cdot ||$ is the vector or scalar L^2 norm and $|| \cdot ||_{-1}$ is meant in the usual sense to be the norm induced by the dual of the H^1 inner product. Now minimization of F makes sense under very general smoothness ($f \in H^{-1}$) and regularity (H^1) assumptions. Unfortunately, this form is too sensitive to variations in a. To see this, first note that the *homogeneous form* $F(\mathbf{u}, p; 0)$ is generally equivalent to the *diagonal form*

$$G(\mathbf{u}, p) \equiv ||\mathbf{u}/\sqrt{a}||^2 + ||\sqrt{a} \nabla p||^2,$$

that is, $F(\mathbf{u}, p; 0) \sim G(\mathbf{u}, p)$. This follows easily in the simplest case $a = 1$ and $\mathbf{b} = \mathbf{0}$, where the Hessians of F and G are as follows:

$$F'' = \begin{pmatrix} I + \nabla L^{-1} \nabla * & -\nabla \\ -\nabla^* & \nabla^* \nabla \end{pmatrix} \text{ and } G'' = \begin{pmatrix} I & O \\ O & \nabla^* \nabla \end{pmatrix}.$$

Here, the representation corresponds to $\begin{pmatrix} \mathbf{u} \\ p \end{pmatrix}$ and the operator L is defined by $L = I + \text{diag}(\nabla^* \nabla)$, which comes from the H^{-1} norm. It is easy to see that the off-diagonal term $-\nabla$ is subdominant because $\mathcal{E} \equiv (I + \nabla L^{-1} \nabla^*)^{-1/2} (-\nabla)(\nabla^* \nabla)^{-1/2}$ satisfies $\mathcal{E}^* \mathcal{E} << I$ (i.e., $||\mathcal{E}|| << 1$). What is important to note is that this is a *scale* bound, not an *order* bound: $\mathcal{E}^* \mathcal{E}$ is dominated by I in scale, not because it is of lower (derivative) order. For the L^2 norm version of FOSLS introduced below, the corresponding scaled off-diagonal term $\mathcal{E}$ is such that $\mathcal{E}^* \mathcal{E}$ is of 'negative order,' meaning in a loose sense that the off-diagonal terms of F'' involve lower derivatives than the diagonal terms. This yields a much stronger sense of diagonal equivalence, which in turn leads to several additional benefits as we suggest below.

Now, in general, the sharpness of the equivalence $F(\mathbf{u}, p; 0) \sim G(\mathbf{u}, p)$ depends intimately on the magnitude $||\sqrt{a}||_\infty$. For the case $a = 0(\frac{1}{\epsilon})$, $F(\mathbf{u}, 0; 0)$ is essentially $|| \nabla^* \mathbf{u}||_{-1}^2 + \epsilon ||\mathbf{u}||^2$. For gradient fields ($\mathbf{u} = \nabla \theta$), we have $F(\mathbf{u}, 0; 0) \sim (1 + \epsilon) ||\mathbf{u}||^2 \sim \frac{1+\epsilon}{\epsilon} G(\mathbf{u}, p)$. For curl fields ($\mathbf{u} = \nabla \times \mathbf{v}$), $F(\mathbf{u}, 0; 0) \sim \epsilon ||\mathbf{u}||^2 \sim G(\mathbf{u}, p)$. Thus, the equivalence of F to G must feel swings of at least $\frac{1+\epsilon}{\epsilon} = 1 + 0(||a||_\infty)$. This dependence can lead to degradation of discretization and algebraic solution techniques.

A choice that appears to avoid this scaling trouble is

$$F(\mathbf{u},p;f) \equiv ||\frac{1}{\sqrt{a}}(\nabla^* \mathbf{u} + \mathbf{b}\cdot(\mathbf{u}/a) - f)||_{-1}^2 + ||\mathbf{u}/\sqrt{a} - \sqrt{a}\nabla p||^2.$$

Generally, we can now expect $F(\mathbf{u},p;0) \sim G(\mathbf{u},p)$, where this diagonal equivalence may be uniform in variations in a. However, there is no apparent way to insulate this equivalence from the deleterious effects of large $||\mathbf{b}||$. This limitation stems from the delicate balance needed between the two terms defining F: scaling up the divergence (first) term by a large coefficient would lead to an unfortunate dominance of this singular term, causing trouble analogous to the large $||\sqrt{a}||_\infty$ case, that weakens the decoupling between the components of $\mathbf{u}$, while scaling up the flux (second) term would severely weaken the decoupling between $\mathbf{u}$ and p. Either scaling therefore weakens the diagonal form equivalence. Other aspects of this weaker sense of diagonal equivalence are that care must be taken: in approximating the H^{-1} norm (i.e., in discretizing L^{-1}); in discretizing the individual variables (loss of approximation properties of p in the $||\sqrt{a}\nabla p||$ sense would directly degrade the approximation properties of $\mathbf{u}$ in the $||\mathbf{u}/\sqrt{a}||$ sense); and in constructing the appropriate algebraic solver.

Generally, the only additional requirement for the L^2 norm version of FOSLS to be applied is that f be in L^2. The added advantages of this version usually make it the method of choice when this is the case. Thus, *with no additional regularity*, we can redefine the functional as

$$F(\mathbf{u},p;f) \equiv ||a\nabla\times(\mathbf{u}/a)||^2 + ||\nabla^*\mathbf{u} + \mathbf{b}\cdot\mathbf{u} - f||^2 + ||\mathbf{u}/\sqrt{a} - \sqrt{a}\nabla p||^2$$

and assert the diagonal form equivalence

$$F(\mathbf{u},p;0) \sim G(\mathbf{u},p) \equiv ||a\nabla\times(\mathbf{u}/a)||^2 + ||\nabla^*\mathbf{u}||^2 + ||\mathbf{u}/\sqrt{a}||^2 + ||\sqrt{a}\nabla p||^2,$$

which is generally uniform in variations in a. With the further assumption of H^2 regularity of the original elliptic equation *without convection* ($\mathbf{b} = \mathbf{0}$), we obtain the stronger vector H^1 uniform equivalence

$$F(\mathbf{u},p;0) \sim G(\mathbf{u},p) \equiv ||\mathrm{diag}(\nabla^*\nabla)\mathbf{u}||^2 + ||\mathbf{u}/\sqrt{a}||^2 + ||\sqrt{a}\nabla p||^2.$$

To understand this equivalence from a heuristic standpoint, consider the simple case $a = 1$ and $\mathbf{b} = \mathbf{0}$. Then

$$F''(\mathbf{u},p;0) = \begin{pmatrix} I + \mathrm{diag}(\nabla^*\nabla) & -\nabla \\ -\nabla^* & \nabla^*\nabla \end{pmatrix} \text{ and } G''(\mathbf{u},p) = \begin{pmatrix} I + \mathrm{diag}(\nabla^*\nabla) & 0 \\ 0 & \nabla^*\nabla \end{pmatrix}.$$

Now the variable coupling is dictated by the scaled off-diagonal term $\mathcal{E} = (I + \mathrm{diag}(\nabla^*\nabla))^{1/2}(-\nabla)(\nabla^*\nabla)^{-1/2}$, which is easily seen to satisfy $\mathcal{E}^*\mathcal{E} << I$ because of scale *and* order ($\mathcal{E}^*\mathcal{E}$ is of 'negative order'). This low order coupling leads to much more flexibility in terms of fast solvers: for example, for small h, the only coupling between variables is via oscillatory components. It also allows for more leeway in individual approximation of variables: for example, $\mathbf{u}$ can be approximate on a grid of mesh size the square root of that for p without adverse effects on the approximation for p. Finally, it enables separate scaling of the individual terms in the functional.

Concerning flexibility in functional scaling, consider the following *two-stage* FOSLS scheme: start by minimizing the functional obtained by deleting the flux term in F (this yields the first-stage functional $||a \bigtriangledown \times(\mathbf{u}/a)||^2 + || \bigtriangledown^* \mathbf{u} + \mathbf{b} - \mathbf{u} \cdot f||^2$ that is well posed in $\mathbf{u}$ alone), then recover p by fixing $\mathbf{u}$ to be the resulting approximation and minimizing the flux term $(||\mathbf{u}/\sqrt{a} - \sqrt{a} \bigtriangledown p||^2)$ alone for p. It is easy to see that this two-stage scheme can usually be done with no real degradation in accuracy or efficiency. However, a more important issue here is the implication that this scale flexibility has on designing functionals for large Reynolds number flow (i.e., large $||\mathbf{b}||$). To see this implication in its simplest setting, consider the model case

$$\bigtriangledown^* \bigtriangledown p + bp_x = f$$

We can reduce this equation by an exponential transformation to the diffusion-only-type case

$$-e^{bx}(e^{-bx}p_x)_x - p_{yy} = f.$$

Here we assume that $x \in [0, 1)$ and $b \geq 0$ so that $e^{-bx} \in (0, 1)$. With this important *damping* scale term in mind, then an appropriate FOSLS functional can be obtained in a straightforward way, but by carefully ensuring that the flux terms do not dominate the div-curl (first and second) terms. This leads to
$F(u, v, p; f) \equiv$

$$\int\int [e^{-bx}((e^{bx}u - p_x)^2 + (v - p_y)^2) + (e^{\frac{bx}{2}}u_x + e^{-\frac{bx}{2}}v_y + e^{-\frac{bx}{2}}f)^2 + (e^{\frac{bx}{2}}u_y - e^{-\frac{bx}{2}}v_x)^2]dxdy.$$

It is better now to absorb e^{bx} into u (replace $e^{bx}u$ by u), so that we can in essence extract out the original transformation, with the aim of making the resulting functional less sensitive to the precise form of the exponential transformation:
$F(u, v, p; f) =$

$$\int\int [(u - p_x)^2 + (v - p_y)^2 + (u_x - bu + v_y + f)^2 + (u_y - v_x)^2]e^{-bx}dxdy.$$

What we have done is: use a *temporary* transformation of scale to recast the original (convection-diffusion) problem to familiar (diffusion) form; then apply FOSLS while exploiting the flexibility in scale to maintain diagonal equivalence; then transform back so that the scale appears only as an integral weighting. This approach is important because it ensures, under fairly general conditions, that the functional is uniformly equivalent to its diagonal form:

$$F(u, v, p; 0) \sim \int\int [u^2 + (u_x + bu)^2 + u_y^2 + v^2 + v_x^2 + v_y^2 + p_x^2 + p_y^2]e^{-bx}dxdy.$$

What is essential here is that the equivalence holds equally well for arbitrarily large Reynolds number b.

Note that the functional obtained by this approach represents a simple modification of the basic FOSLS functional

$$|| \bigtriangledown \times \mathbf{u}||^2 + || \bigtriangledown^* \mathbf{u} + \mathbf{b}^* \mathbf{u} - f||^2 + ||\mathbf{u} - \bigtriangledown p||^2.$$

The only difference is a weighting of the L^2 norm by a term that attenuates exponentially in the streamwise direction into what is typically the boundary layer. This amounts to damping the influence of errors in all variables in the boundary layer, which seems appropriate because it has the effect of protecting the rest of the region (free stream) from contamination caused by the general inability to approximate the rapidly changing variables there. The functional equivalence to the H^1 -like norm weighted in this way ensures that free-stream discretization accuracy is optimal in all variables in the unscaled H^1 sense, and that multigrid converges with the usual elliptic speed.

Convection-dominated applications are extremely challenging and subtle in all but the simplest cases, such as that considered here. It is therefore difficult to know if the simple approach developed above for constant- coefficient problems can be extended to those of more practical interest. How general this approach is and how sensitive it is to the proper choice of scale remains an open question. However, it seems likely that a reasonable and efficient strategy for local approximation of the weighting (based on evolving elementwise estimates of the stream direction) could be devised for a fairly general class of variable-coefficient problems. In fact, even nonlinear convection-diffusion problems may be amenable to this approach, especially if full multigrid schemes are used to obtain good *initial* estimates from coarser grids.

Note that uniform diagonal equivalence means that the overall accuracy and efficiency can be considered separately for each variable. Since the variable coupling is subdominant in scale and order, then *relative accuracy within each variable is assured in its associated scalar norm.*

Two additional benefits come with the least-squares approach: sharp a posteriori error estimates, enabled because the true minimum value of the functional is known (zero), and natural treatment nonconforming finite elements, enabled because the functional is directly posed in weak form. For either version of FOSLS, the quantities

$$F(\mathbf{v}, q; f) \qquad \text{and} \qquad \frac{F(\mathbf{v}, q; f)}{F(\mathbf{0}, 0; f)}$$

serve as *precise* measures of the absolute and relative errors in the approximation $(\mathbf{v}, q)$, no matter how it was obtained. Although this assumes that $F(\cdot, \cdot; 0)$ itself is acceptable as an error norm, in any case these quantities provide *sharp* error estimates in terms of the appropriate Sobolev norm on $(\mathbf{v}, q)$, depending on the norm equivalence established for $F(\cdot, \cdot; 0)$. The advantages in terms of nonconforming elements is illustrated by considering a space that violates whatever boundary conditions might be present. For example, for the convection-diffusion case considered above, the boundary condition $p = g$ would naturally be replaced by $p = g^h$, where g^h is the trace of some function in the discrete space for p. (A similar approximation for FOSLS would be made for $\mathbf{u}$ components as well.) Of course, any finite element method must account for the error induced directly by this approximation, but, for standard finite elements applied to the original scalar convection-diffusion problem, we would also have to account for the error in the weak form caused by the new boundary integral terms that this approximation generates. However, no such terms are generated in the FOSLS functional since the weak form occurs directly, without integration by parts, so the only issue here is how well the nonconforming subspace approximates the original solution. In words borrowed from Gil Strang, no variational crimes have been committed. Thus, said in finite element terms, Strang's lemma reduces to the much simpler Ceá's lemma for FOSLS functionals.

Further Information

See the papers (and the references that they cite) available by anonymous ftp from

amath.colorado.edu/pub/fosls

or by way of the World Wide Web from

http://amath-www.colorado.edu/appm/faculty/stevem/Home.html

where a compressed postscript version of this document (FOSLoSophy.ps.Z) also resides.