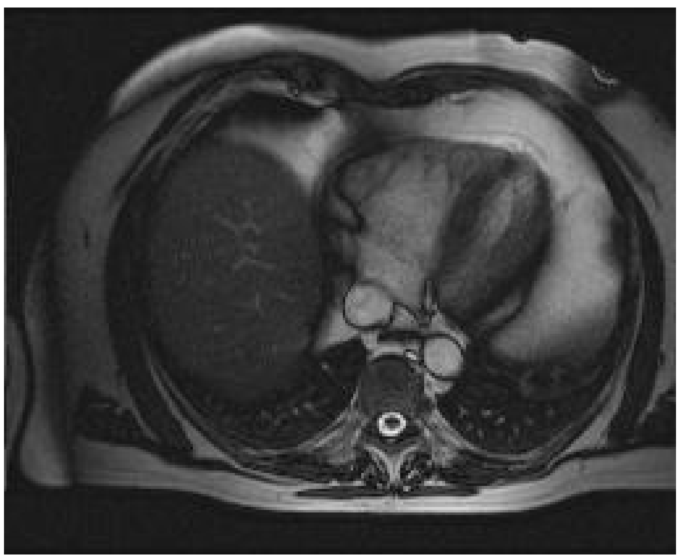
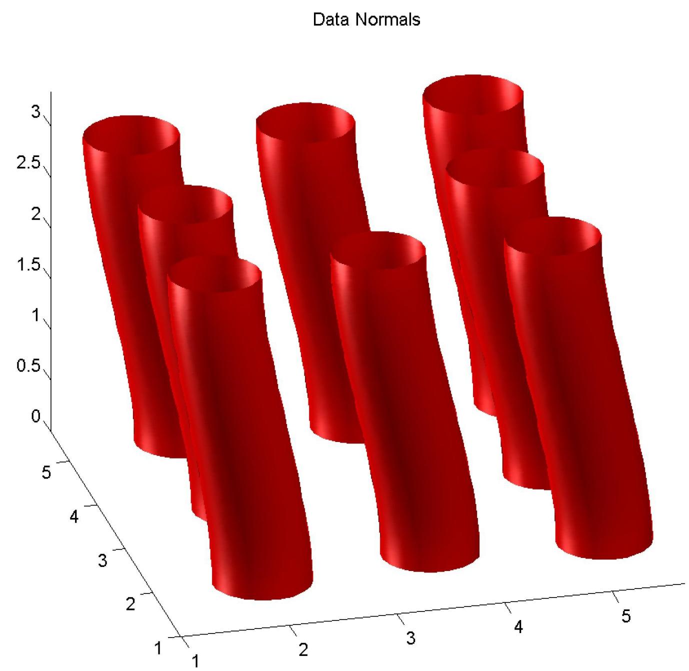
Sobolev gradients and negative norms for image decomposition.

Walter Richardson, Jr.
University of Texas at San Antonio
walter.richardson@utsa.edu





Least Squares Prototype

Steady-state diffusive transport:

$$-\boldsymbol{\nabla} \cdot (\boldsymbol{A}\boldsymbol{\nabla} u) + cu = f, \qquad x \in \Omega$$

Mixed Method: Second-order equation $\Rightarrow$ first-order system using flux

$$\begin{aligned} \boldsymbol{\sigma} \ + \ \boldsymbol{A}\boldsymbol{\nabla} u &= 0 \\ \boldsymbol{\nabla} \cdot \sigma \ + \ cu &= f \end{aligned}$$

Boundary conditions $u|_{\partial\Omega} = 0, \quad \boldsymbol{\sigma} \cdot n = 0$; variational formulation is to minimize the norms of the residuals

$$J(u, \boldsymbol{\sigma}) = \frac{1}{2} \int_\Omega |\boldsymbol{\nabla} \cdot \boldsymbol{\sigma} \ + \ cu - f|^2 \quad + \quad \frac{1}{2} \int_\Omega \|\boldsymbol{\sigma} + \boldsymbol{A}\boldsymbol{\nabla} u\|^2$$

Least Squares Pros & Cons

1. Advantages:

   (a) generality & convenience
   (b) symmetric form even for non-self adjoint PDE's
   (c) less sensitive to changes in PDE type (transonic flow)
   (d) ease of error evaluation
   (e) mixed least squares finite elements without restrictive inf-sup condition of Galerkin mixed methods

2. Disadvantages:

   (a) ill-conditioning - analogous to solving normal equations for $Au = b$
   (b) degradation of iterative convergence
   (c) subtle scaling issues

Quadratic Functionals Different Norms

Standard LSFEM uses functional

$$Q_1(\delta, v) = \frac{1}{2} \|\boldsymbol{\nabla} \cdot \delta + \chi v - f\|_{L^2}^2 + \frac{1}{2} \|\mathcal{A}^{-\frac{1}{2}}(\delta + \mathcal{A}\boldsymbol{\nabla} v)\|_{L^2}^2$$

Too many derivatives on σ, L^2 norm too strong in first term above: Try a weaker norm.

$$Q_2(\delta, v) = \frac{1}{2} \|\boldsymbol{\nabla} \cdot \delta + \chi v - f\|_{H^{-1}}^2 + \frac{1}{2} \|\mathcal{A}^{-\frac{1}{2}}(\delta + \mathcal{A}\boldsymbol{\nabla} v)\|_{L^2}^2$$

If $s = \min(k, r)$ where C^0 FE spaces of degree k and r are used for u_h,σ_h:

$$\|e_u\|_m + \|e_\sigma\|_m \leq Ch^{s+1-m} \left(\|u\|_{s+1} + \|\sigma\|_{s+1}\right), \qquad m = 0, 1$$

Continuous Gradient Descent

Find zeros of differential operator $F : \{H, \langle \cdot, \cdot \rangle\} \longrightarrow \{K, (\cdot, \cdot)\}$. Different inner products in domain space give different descent directions. If $J(u) = \frac{1}{2}\|F(u)\|^2$, gradient is $\boldsymbol{\nabla} J(u) = F'(u)^\star F(u)$ Where $'$ denotes Fréchet derivative, * denotes adjoint Find $z : [0, \infty) \to H$ satisfying

$$\frac{dz}{dt} = -\boldsymbol{\nabla} J(z(t)) \qquad\qquad z(0) = z_0$$

Sufficient conditions for convergence:

1. Convexity: $\langle J''(u)v, v\rangle \geq m\|v\|^2$
2. Gradient Inequality: $\|\boldsymbol{\nabla} J(u)\| \equiv \|F'(u)^\star F(u)\| \geq C\|F(u)\|$

Euclidean vs. Sobolev Gradients

Discretize the elliptic system using finite differences, finite elements, etc. Minimize discrete LS funtional $\tilde{J}(u, \boldsymbol{\sigma})$ using steepest descent with optimum stepsize (for linear PDE) with $\boldsymbol{w} = (u, \boldsymbol{\sigma})$:

$$\boldsymbol{w}_{k+1} = \boldsymbol{w}_k - \alpha_k \boldsymbol{\nabla} \tilde{J}(\boldsymbol{w}_k) \qquad \alpha_k = \|\boldsymbol{\nabla} J(\boldsymbol{w}_k)\|^2 / \|\boldsymbol{A} \boldsymbol{\nabla} J(\boldsymbol{w}_k)\|^2$$

Euclidean gradient on both components: Slow convergence, rough gradient. Sobolev descent: smooth gradients, rapid convergence.

$$\langle u, v \rangle = (u, v) + (Du, Dv) = ([I + D^\star D]u, v)$$

Component u lives in H^1, do descent in that space use Riesz represenation to obtain Soboleve gradient: Sobolev gradient obtained by solving linear system $(I + D^\star D)\boldsymbol{\nabla}_S J(x) = \boldsymbol{\nabla} J(x)$

Descent as General Iterative Method

General iterative method for $Au = b$:

$$u_{n+1} \;=\; u_n - \boldsymbol{H}_n(\boldsymbol{A}u_n - b)$$

Steepest descent applied to $J(u) = \frac{1}{2}\langle \boldsymbol{A}u, u\rangle - \langle \boldsymbol{b}, u\rangle + C$ uses $\boldsymbol{H}_n = \beta_n \boldsymbol{I}$ where β_n chosen optimally. Upper bound on convergence - Kantorovich's inequality:

$$\frac{J(u_n + \beta_n r_n) - J(u_\infty)}{\sigma(u_n) - \sigma(u_\infty)} \;\leq\; \left(\frac{M - m}{M + m}\right)^2$$

where $M = \lambda_{max}$ and $m = \lambda_{min}$ Sobolev gradient preconditions by $\boldsymbol{H} = (I + D^\star D)^{-1}$.

Sobolev Gradients and H^{-1}

Given $f \in L_2$, $\ell(\cdot) = \langle f, \cdot \rangle_0$ defines bounded linear functional on H^1,let w represent ℓ

$$(f, x) = \langle w, x \rangle_1 = (w, x) + (\nabla w, \nabla x) = (w - w'', x)$$

Recall H^{-1} denotes dual of H^1; how to compute norm?

$$\|f\|_{-1}^2 = \sup_{x \in H^1} \frac{(f, x)^2}{\|x\|_1^2} = \sup_{x \in H^1} \frac{\langle w, x \rangle_1^2}{\|x\|_1^2} = \|w\|_1^2 = (w, f) = (Tf, f)$$

where $T : L^2 \to H^1$ is defined by $(I - D^2)w = f$. More generally, $T \Rightarrow T_h = (I - \Delta_h)^{-1}$ which is precisely the $(I + D^\star D)$ relating Sobolev and Euclidean gradients.

Digital Image Reconstruction

Consider splitting (Meyer) an image f up into components $f = u + v$, where u is a smoothed version of f and v represents texture+noise. A variational formulation for the image model is

$$J(u, \lambda) = \frac{1}{p}\|\nabla u\|^p + \lambda\ \frac{1}{q}\|v\|^q$$

$p = 2, q = 2$ Heat equation
$p = 1, q = 2$ Rudin-Osher-Fatemi $TV - L^2$ model
$p = 1, q = 1$ Chan-Vese $TV - L^1$ preserves contrast & geometry

LSFEM with Sobolev gradients would suggest negative, even fractional, norms useful. Connection to real interpolation spaces via K-functional:
$K(t, u) = \inf_{v \in X} \left(\|u - v\|_Y^2\ +\ t^2\|v\|_X^2\right)^{\frac{1}{2}}$

Helmholtz Equation

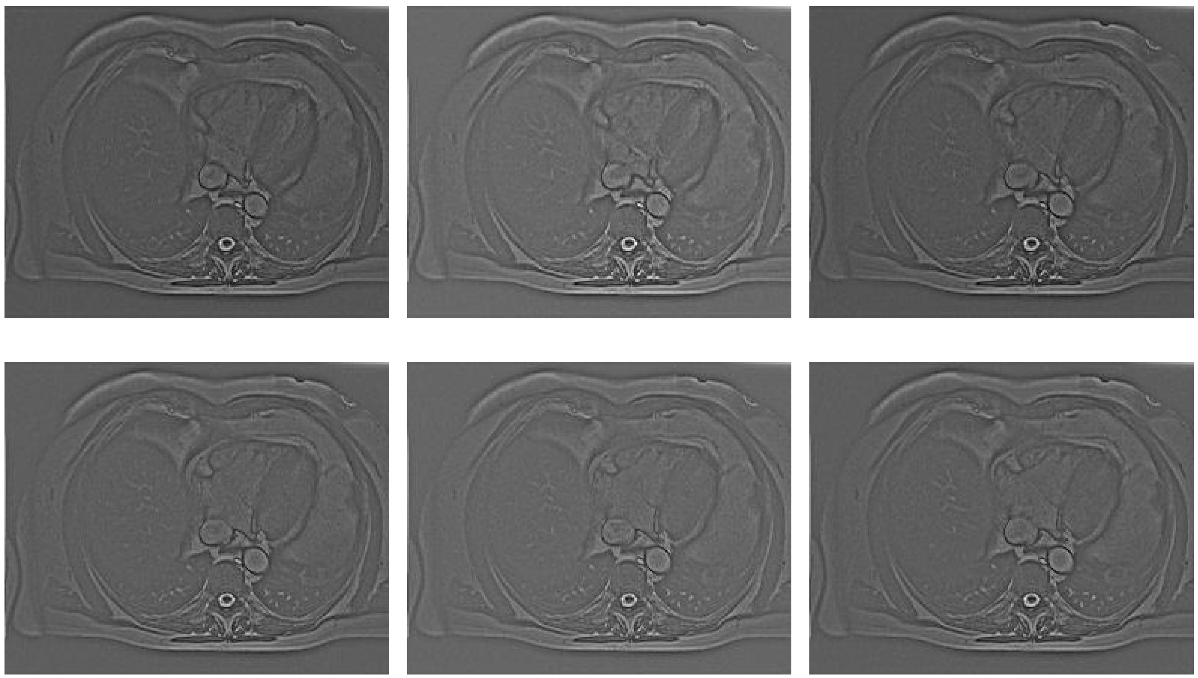
Decompose $f = u + v$, where $u \in H^1$ and $v \in L^2$. Minimize

$$\min_{u \in H^1} J(u) = \|u\|^2_{H^1} \; + \; \lambda \|f - u\|^2_{H^0}$$

with zero Neumann boundary conditions as most reasonable.
$J'(u)h = 2(u,h) + 2(\nabla u, \nabla h) + 2\lambda(f-u, -h)$, Euler-Lagrange equation is $\Delta u - (1+\lambda)u \; = \; -\lambda f$. Below is the $f = u + v$ decomposition.

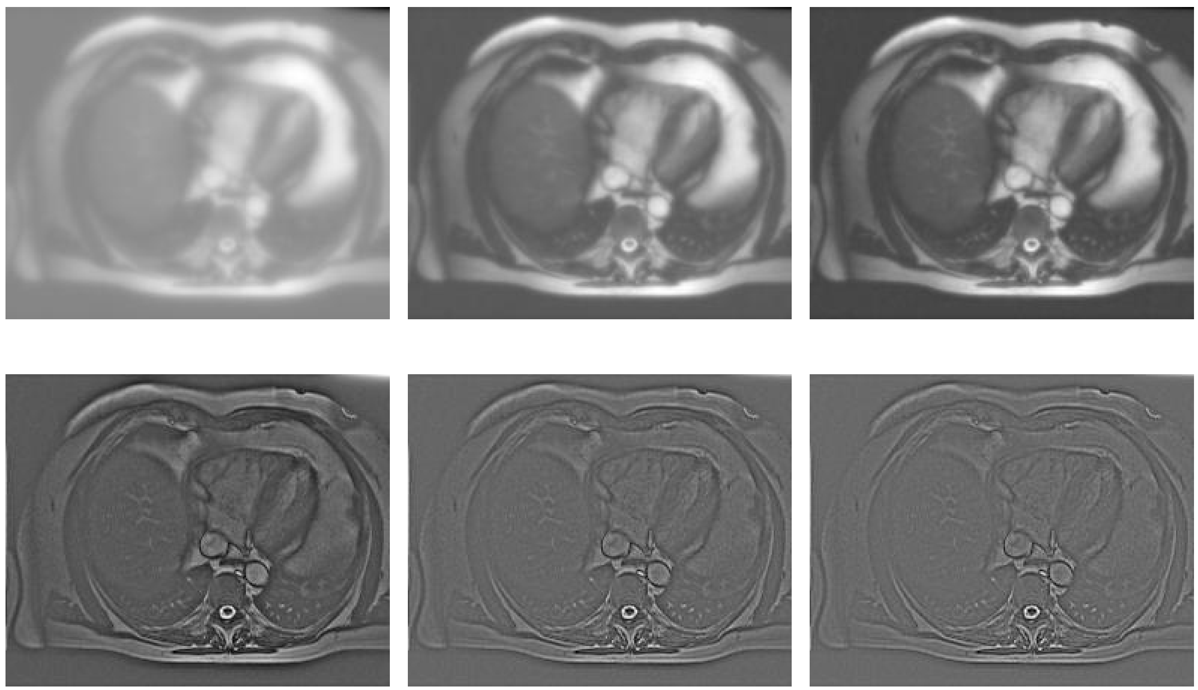
Fast Solvers for 3D images

To be useful, the deomposition/preconditioning must be done quickly. Critical for 3D images or 2D time sequences in which processing should be done more quickly than video capture. Use fast Poisson solver, multigrid/multipole methods, wavelets, or adaptive grids.



Varying the Scale Parameter λ

Weighting parameter λ controls which details are to be removed from the smoothed image. Below are u, v decompositions for $\lambda = 0.1, 0.4, 0.7$.

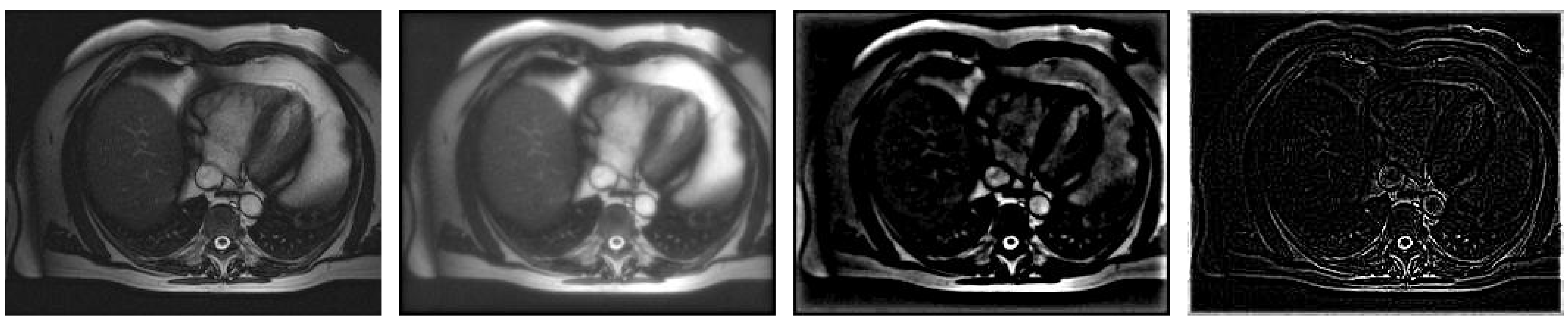


Negative Norm Decomposition

Decompose $f = u + v + w$, where $u \in H^1$,$v \in H^{-1}$, $w \in L^2$. Minimize

$$\min_{u \in H^1} J(u,v) = \|u\|_{H^1}^2 + \lambda_1 \|v\|_{H^{-1}}^2 + \lambda_2 \|f - (u+v)\|_{H^0}^2$$

Below original image at left, followed by u, v, w decomposition.



References

[1] J. W. Neuberger, *Sobolev Gradients and Differential Equations*, Springer-Verlag, Berlin, 1997.

[2] W. B. Richardson, Jr. "Sobolev Gradient Preconditioning for PDE Applications", *Iterative Methods in Scientific Computation IV, IMACS Series in Computational and Applied Mathematics*, v. 5, p. 223-234, 1999.

[3] W. B. Richardson, Jr., "Steepest Descent using Smooth Gradients", *Journal of Applied Mathematics and Computation*, v. 112, p. 241-254, 2000.

[4] G. F. Carey and W. B. Richardson, "A note on least-squares methods," to appear in *Communciations in Numerical Methods in Engineering*.

[5] W. B. Richardson, "High Order Sobolev Preconditioning", to appear in *Nonlinear Analysis*.