On Jacobian model reduction for the time-domain inverse problem

V. Simoncini

Dipartimento di Matematica, Università di Bologna

valeria@dm.unibo.it

joint work with V. Druskin & M. Zaslavsky (Schlumberger-Doll Res.)

The parameter identification problem

Diffusive Maxwell's system:

$$\nabla \times \rho \nabla \times H + H_t = 0, \qquad H(0) = \varphi, \quad t > 0$$

$H(t)$: vector-magnetic field in $\mathbb{R}^3 \times [0, \infty]$

$\rho \in L_\infty[\mathbb{R}^3]$: uniformly positive electrical resistivity distribution

φ: compactly supported magnetic source distribution

$\star$ Other problems lead to similar setting (e.g., divergence op.)

Aim: Recover an approximate $\rho = \rho(x)$ from observed data

Working hypothesis: $\rho \in \rho_0 + \mathcal{S}$, with $\mathcal{S}$ compact set

This talk is about a proof of concept $\Rightarrow$ 1D problem

The semidiscrete parabolic problem

By, e.g., FD approximation we obtain $\qquad n$ large

$$A(r)u + u_t = 0, \qquad u(0) = b \in \mathbb{R}^n, \quad 0 < A(r) = A^T(r) \in \mathbb{R}^{n\times n}$$

$r \in r_0 + S$, $\qquad S$: space of admissible functions

$$A(r) = D^T \mathrm{diag}(r) D, \qquad D \approx \nabla\times, \text{ or } D \approx \nabla$$

Analytic solution : $\qquad u(t,r) = e^{-tA(r)}b$

The semidiscrete parabolic problem

By, e.g., FD approximation we obtain n large

$$A(r)u + u_t = 0, \qquad u(0) = b \in \mathbb{R}^n, \quad 0 < A(r) = A^T(r) \in \mathbb{R}^{n\times n}$$

$r \in r_0 + S$, S: space of admissible functions

$$A(r) = D^T \mathrm{diag}(r) D, \qquad D \approx \nabla\times, \text{ or } D \approx \nabla$$

Analytic solution : $u(t,r) = e^{-tA(r)} b$

For a given r, operator associated with the forward problem:

$$y(t,r) = b^* e^{-tA(r)} b \ \in \ L_2[0,\infty]$$ Single Input = Single Output

to be compared w/ observable measurements:

$$d(t) = y(t, r_{\mathrm{true}}) + \Delta(t)$$

$\Delta(t)$: measurement error r_{true}: sought after resistivity

The model for the numerical solution

Assuming $r_{\text{true}} \in r_0 + S$, we state the problem as:

$$r_{\text{approx}} = \arg \min_{r \in r_0 + S} \frac{1}{2} \|y(\cdot, r) - d(\cdot)\|^2_{L_2}$$

Remark: For stability reasons, regularization is usually employed:

$$r_{\text{approx}} = \arg \min_{r \in r_0 + S} \frac{1}{2} \left(\|y(\cdot, r) - d(\cdot)\|^2_{L_2} + \lambda \Phi^d(r) \right)$$

Proof of concept $\Rightarrow$ no regularization, $\lambda = 0$

Gauss-Newton iteration and the Jacobian

For the forward operator $y(t,r) = b^* e^{-tA(r)} b$,

$$\delta y = J \delta r, \qquad \text{where} \quad \delta y = \frac{d}{d\sigma} y(t, r + \sigma \delta r)|_{\sigma=0}$$

$$J : \delta r \in \mathbb{R}^k \mapsto \delta y \in L_2[0,\infty]$$

Gauss-Newton iteration:

$$r_{j+1} = r_j - (J^T J)^{-1} J^T \diamond (y(t, r_j) - d(t))$$

($\diamond$: inner product in L_2)

Problem: Computation of the Jacobian is highly expensive

$$\Downarrow$$

Approximate Jacobian J_m

Towards Jacobian approximation

Let $\mathcal{U}_m$ be m-dim subspace, $m \leq n$, (θ_j, z_j) Ritz pairs of $A(r)$ in $\mathcal{U}_m$

Approximation of y in $\mathcal{U}_m$:

$$y_m(t,r) = b^T u_m(t,r) = \sum_{i=1}^{m} e^{-\theta_i t} c_i, \qquad c_i = (z_i^T b)^2$$

and

$$\delta y_m(t, \delta r) = \frac{d}{ds} y_m(r + s\delta r)|_{s=0} = \sum_{i=1}^{m} e^{-\theta_i t} \delta c_i - \sum_{i=1}^{m} t e^{-\theta_i t} c_i \delta\theta_i$$

$(\delta c_i = \frac{d}{ds} c_i, \quad \delta\theta_i = \frac{d}{ds}\theta_i, \quad i = 1, \ldots, m)$

Which $\mathcal{U}_m$? best choice minimizes the “error” $y \to y_m$

$$\delta y_m(t,\delta r) = \frac{d}{ds} y_m(r+s\delta r)|_{s=0} = \sum_{i=1}^{m} e^{-\theta_i t}\delta c_i - \sum_{i=1}^{m} t e^{-\theta_i t} c_i \delta\theta_i$$

that is,

$$\delta y_m(t,\delta r) = \underbrace{[e^{-t\theta_1},\ldots,e^{-t\theta_m},te^{-t\theta_1},\ldots,te^{-t\theta_m}]}_{\mathbf{e}(t)^T} \begin{bmatrix} \delta c_1 \\ \vdots \\ \delta c_m \\ -c_1\delta\theta_1 \\ \vdots \\ -c_m\delta\theta_n \end{bmatrix}$$

The red part implicitly defines a matrix B_m, so that

$$\delta y_m(t,\delta r) = \mathbf{e}_m(t)^T B_m \delta r \qquad \Rightarrow \qquad J_m = \mathbf{e}_m(t)^T B_m$$

The Jacobian $J_m = \mathbf{e}_m(t)^T B_m$

Necessary condition: r satisfies $\quad J_m^T \diamond (y_m(\cdot, r) - d) = 0$

The Jacobian $J_m = \mathbf{e}_m(t)^T B_m$

Necessary condition: r satisfies $\quad J_m^T \diamond (y_m(\cdot, r) - d) = 0$

TH: Let $d = y(r_{\rm true})$, and y_m be obtained in

$\mathcal{U}_m = \text{span}\{(A + s_1 I)^{-1} b, \ldots, (A + s_m I)^{-1} b\}$

If $s_j = \theta_j, \quad j = 1, \ldots, m$, then

$$J_m^T \diamond (y_m(\cdot, r_{\rm true}) - d(\cdot)) = 0$$

(using [Meier & Luenberger, '67] and H_2-optimality of $\mathcal{U}_m$)

$\Rightarrow$ If $r_\infty \in r_0 + S$ is the unique solution in $\mathcal{U}_m$ to $J_m^T \diamond (y_m(\cdot, r) - d) = 0$ with $d(t) = y(r_{\rm true}, t)$, then $r_\infty = r_{\rm true}$ iff $s_j = \theta_j, \forall j$

The Gauss-Newton iteration

$$r_{j+1} = r_j - (J_m^T J_m)^{-1} J_m^T \diamond (y_m(t, r_j) - d(t)), \qquad J_m = \mathbf{e}_m(t) B_m(r)$$

⋆ Computing $J_m^T J_m$: $\quad J_m^T J_m = B^T \mathbf{e}_m(t) \diamond \mathbf{e}_m(t)^T B$ with

$$\begin{aligned} \mathbf{e}_m(t) \diamond \mathbf{e}_m(t)^T &= \begin{bmatrix} e^{-t\theta_i} \diamond e^{-t\theta_j} & e^{-t\theta_i} \diamond te^{-t\theta_j} \\ te^{-t\theta_i} \diamond e^{-t\theta_j} & te^{-t\theta_i} \diamond te^{-t\theta_j} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{\theta_i+\theta_j} & \frac{1}{(\theta_i+\theta_j)^2} \\ \frac{1}{(\theta_i+\theta_j)^2} & \frac{2}{(\theta_i+\theta_j)^3} \end{bmatrix} \in \mathbb{R}^{2m\times 2m} \end{aligned}$$

⋆ Computing $J_m^T \diamond (y_m(t, r_j) - d(t))$: $\quad J_m^T f(t) = B^T(\mathbf{e}_m(t) \diamond f(t))$
quadrature rule

The Gauss-Newton iteration. cont'd

$$r_{j+1} = r_j - (J_m^T J_m)^{-1} J_m^T \diamond (y_m(t, r_j) - d(t)), \qquad J_m = \mathbf{e}_m(t) B_m(r)$$

The matrix B_m (exact for $m = n$):

$$B_m \approx -[2Q_m \circ W_m,\, Q_m \circ Q_m]^T$$

$$Q_m = D[z_1(z_1^T b), \ldots, z_m(z_m^T b)]$$
$$W_m = DU_m[(H_m - \theta_1 I)^\dagger U_m^T b, \ldots, (H_m - \theta_m I)^\dagger U_m^T b]$$

with $H_m = U_m^T A(r) U_m$

(θ_j, z_j) Ritz pairs of $A(r)$ in $\mathcal{U}_m = \mathrm{Range}(U_m)$

for $m \ll n$ inexpensive computation !

The practicalities that make things work

Given S (admissible function space)

$$r_{j+1} = r_j - (J_m^T J_m)^{-1} J_m^T \diamond (y_m(t, r_j) - d(t)), \qquad J_m = \mathbf{e}_m(t) B_m(r)$$

- Chose m small (but with $2m$ larger than k=dim(S))
- Project Gauss-Newton recurrence onto S
- Determine H_2 optimal $\mathcal{U}_m$ for current $A(r_j)$ to compute y_m
  (IRKA, Gugercin etal, '08)

Overall nonlinear recurrence

Given r_0, $A(r_0)$, b, d, $m \ll n$, admissible function space S

for $j = 0, 1, \ldots$

1. Compute Orthonormal basis U_m of H_2-*optimal* $\mathcal{U}_m$ for $A(r_j)$
2. Compute $H_m = U_m^T A(r_j) U_m$ and (θ_i, z_i) Ritz pairs
3. Compute matrix B_m and project onto S
4. Compute matrix $J_m^T J_m = B_m^T (\mathbf{e}_m \diamond \mathbf{e}_m^T) B_m$
5. Update $r_{j+1} = r_j - (J_m^T J_m)^{-1} (B_m^T \mathbf{e}_m \diamond (y_m(\cdot, r_j) - d))$
6. If converged, then stop

Note: For m sufficiently small, $J_m^T J_m$ well-conditioned

Numerical experiments

$$A = D^T \mathrm{diag}(r) D \in \mathbb{R}^{n\times n}, \quad D = (n-1)\begin{bmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & 1 & -1 & \\ & & & \ddots & \ddots \\ & & & & 1 \end{bmatrix}$$

$n = 900$, $b = 1$ $\qquad$ $r_{\mathrm{true}} \in r_0 + S$, $\quad$ r_0 constant

dim(S)=1 (parametric inversion), no noise

$d(t) = y(r_{true}, t), \quad t \geq 0 \quad S = [1, \ldots, 1, 0, \ldots, 0]^T$

m	G-N it	$\|\|y_m - d\|\|_{L_2}$	$\|\|r_j - r_{\mathrm{true}}\|\|$	$\|\|J_m^T(y_m - d)\|\|$	# IRKA its
2	1	2.5235e-03	2.1208e+01	3.0352e-07	19
	2	6.4960e-04	1.7030e+00	1.4994e-08	14
	3	6.2789e-04	1.3646e-01	1.1522e-09	13
	4	6.2776e-04	1.2988e-02	1.1003e-10	11
	5	6.2775e-04	1.2204e-03	1.0334e-11	10
	6	6.2775e-04	1.1465e-04	9.7212e-13	8
	7	6.2775e-04	1.0934e-05	9.1438e-14	7
	8	6.2775e-04	8.7870e-07	8.6042e-15	5
8	1	2.4410e-03	2.1208e+01	3.3444e-07	61
	2	3.3750e-04	3.4000e+00	3.4391e-08	53
	3	8.5406e-06	8.8060e-02	8.2569e-10	48
	4	5.4750e-07	8.2302e-05	7.7139e-13	39
	5	5.4744e-07	1.0997e-07	2.1110e-16	21

dim(S)=1, noisy data: $d(t) = y(r_{true}, t) + \Delta(t) \quad t \geq 0$

$S = [1, \ldots, 1, 0, \ldots, 0]^T$, $\quad m = 2 \quad \|y(r_{true}, t)\|_{L_2} \approx 0.16$

$\|\Delta\|_{L_2}$	GN it	$\|\|y_m - d\|\|_{L_2}$	$\|\|r_j - r_{\rm true}\|\|$	$\|\|J_m^T(y_m - d)\|\|$	# IRKA its
1e-03	1	7.5218e-03	1.0606e+02	9.2578e-07	19
	2	2.3974e-03	4.6569e+01	1.1219e-07	14
	3	1.1426e-03	6.7352e+00	8.3070e-09	14
	4	1.1179e-03	4.4495e-02	4.4094e-10	12
	5	1.1178e-03	3.6355e-01	4.0665e-11	11
	6	1.1178e-03	3.2620e-01	3.6458e-12	9
1e-05	1	7.4695e-03	1.0606e+02	9.2743e-07	19
	2	2.1866e-03	4.6463e+01	1.1238e-07	15
	3	5.5489e-04	6.4620e+00	8.3449e-09	14
	4	5.0042e-04	3.8368e-01	4.4169e-10	12
	5	5.0023e-04	2.7590e-02	4.0805e-11	11
	6	5.0023e-04	1.0122e-02	3.6642e-12	9
	7	5.0023e-04	6.7332e-03	3.2991e-13	8

dim(S)=1, noisy data: $d(t) = y(r_{true}, t) + \Delta(t) \quad t \geq 0$

$S = [1, \ldots, 1, 0, \ldots, 0]^T$, $m = 8$ $\|y(r_{true}, t)\|_{L_2} \approx 0.16$

$\|\Delta\|_{L_2}$	GN it	$\|\|y_m - d\|\|_{L_2}$	$\|\|r_j - r_{\text{true}}\|\|$	$\|\|J_m^T(y_m - d)\|\|$	# IRKA its
1e-03	1	2.6239e-03	2.1208e+01	3.3236e-07	61
	2	1.0541e-03	3.5108e+00	3.4071e-08	51
	3	9.9993e-04	2.4095e-01	8.1221e-10	48
	4	9.9989e-04	1.5480e-01	8.2823e-13	39
1e-05	1	2.4412e-03	2.1208e+01	3.3447e-07	61
	2	3.3770e-04	3.3986e+00	3.4397e-08	51
	3	1.3144e-05	8.6069e-02	8.2613e-10	48
	4	1.0002e-05	1.9618e-03	7.7672e-13	39

dim$(S) > 1$, no noise.

$d(t) = y(r_{true}, t), \;\; t \geq 0$

$S = \mathrm{span}\{x_1, \ldots, x_5\}$ leading eigvecs of $J^T J$

m	G-N its	$\|\|y_m - d\|\|_{L_2}$	$\|\|r_j - r_{\mathrm{true}}\|\|$	$\|\|J_m^T(y_m - d)\|\|$
6	7	3.9625e-06	5.7742e-08	5.1348e-16
12	4	1.8104e-08	3.1126e-08	1.7867e-16
18	4	8.6085e-11	3.0688e-08	9.7134e-16

(corresponding results for noisy data)

dim(S)=1, no noise. Optimality. Final attained accuracy

$d(t) = y(r_{true}, t), \quad t \geq 0 \qquad S = [1, \ldots, 1, 0, \ldots, 0]^T$

m	G-N its	$\|\|y_m - d\|\|_{L_2}$	$\|\|r_j - r_{\text{true}}\|\|$	$\|\|J_m^T(y_m - d)\|\|$	method
m=2	5	6.9103e-04	2.2490e-08	8.8018e-17	IRKA
	6	3.5500e-03	1.9052e+00	4.5795e-17	RKSM
	6	3.5500e-03	1.9052e+00	3.0800e-17	Adapt RKSM
m=4	5	3.6966e-05	2.2826e-08	1.7038e-16	IRKA
	4	5.5125e-05	4.9347e-05	1.9499e-17	RKSM
	4	1.1393e-04	1.1044e-02	1.8129e-16	Adapt RKSM
m=6	4	3.9865e-06	2.2673e-08	1.4685e-16	IRKA
	4	9.0685e-06	4.6856e-08	2.8661e-18	RKSM
	4	4.7859e-06	1.1861e-06	5.8369e-18	Adapt RKSM
m=8	4	5.9811e-07	2.2003e-08	1.2173e-17	IRKA
	4	3.7142e-06	2.1582e-08	2.2253e-17	RKSM
	4	1.1268e-06	4.2661e-08	4.5059e-18	Adapt RKSM

Conclusions

- Promising new conceptual approach
- Perform further numerical experiments
- Develop prototype towards real appl. problem (w/regularization, line search, etc.)

For further info, e–mail: `valeria@dm.unibo.it`