Yin Huang’s Thesis, and Computing Gradients

William Symes
Yin Huang

PhD student in TRIP: 2010.08 - 2016.02

Thesis: *Born Waveform Inversion in Shot Coordinate Domain*

Currently: Amazon, Seattle
Chapter 2

Born waveform inversion via shot record extension, variable projection, differential semblance

[SEG 2015]
Chapter 2

Task: estimate Marmousi from homog. initial guess

Figure 2.2: Marmousi example: Target background model (a) target reflectivity model (b)
Figure 2.3: Marmousi example: Born shot record with index 41.
Figure 2.11: Marmousi example: Inverted reflectivity model at true background model (a); initial background model (b); background model with 18 steps of VPE method (c) and 18 steps of VP method (d) by solving equation 2.9.
Chapter 2

Figure 2.12: Marmousi example: Common image gathers at true background model (a); initial background model (b); background model with 18 steps of VPE method (c) by solving equation 2.9.
Figure 2.13: Marmousi example: Trace comparison of real data (red), predicted data by VPE (green) and VP (blue) methods for far (top) and near (bottom) offsets.
Chapter 2

Bottom line: works, but slow

18 VP its × 50 CG iterations - way too much
Chapter 3

*Flexibly Preconditioned Extended Least Squares Migration in Shot Record Domain*

Joint with Rami Nammour - in review © *Geophysics*
Chapter 3

Task: use ΨDO scaling to precondition inner problem
Chapter 3

ΨDO scaling - Nammour 09, uses Bao-S. 96

Estimate amplitude by 2 Hessian ops

Flexibly Preconditioned CG
Figure 3.15: Marmousi example: convergence curves of numerical methods, (a) normalized data misfit and (b) normalized gradient length.
Figure 3.16: Marmousi example: inverted model perturbation cube after 20 Hessian applications using FPCG (a) and using CG with windowing (b).
Figure 3.20: Marmousi example: data residual, same shot record as in Figure 3.14b, after 20 Hessian applications using FPCG (a), and using CG with windowing (b).
Chapter 3

Bottom line: speedup by factor of 3-4

Much better inner solve with same effort
Chapter 4

Task: evaluate effect of FPCG/CG inner solve on gradient accuracy
Chapter 4

Fast lens over flat reflector

Computed gradient at const background model
Chapter 4

Relative error in

\[
\frac{J[m + h\delta m] - J[m - h\delta m]}{2h}
\]

as approx to \( \langle \nabla J[m], \delta m \rangle \)
Table 4.1: Gradient test at constant background model \( m = (2\text{km/s})^2 \) for different \( dm \) and different numerical methods with 20 applications of LSM Hessian

<table>
<thead>
<tr>
<th>Numerical methods</th>
<th>dm1</th>
<th></th>
<th></th>
<th>dm2</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>h</td>
<td>0.1</td>
<td>0.05</td>
<td>0.025</td>
<td>0.1</td>
<td>0.05</td>
<td>0.025</td>
</tr>
<tr>
<td>FPCG</td>
<td>0.1108</td>
<td>0.0525</td>
<td>0.0643</td>
<td>0.1715</td>
<td>0.0814</td>
<td>0.1346</td>
</tr>
<tr>
<td>CG with ( I_t^2 )</td>
<td>0.1131</td>
<td>0.1147</td>
<td>0.1105</td>
<td>0.0467</td>
<td>0.0579</td>
<td>0.0633</td>
</tr>
<tr>
<td>CG with windowing</td>
<td>0.9890</td>
<td>1.0088</td>
<td>1.0673</td>
<td>1.3502</td>
<td>1.1384</td>
<td>0.6748</td>
</tr>
</tbody>
</table>
Chapter 4

Bottom line: not so hot

Why? Look to nature of tomo op
\[ J[m] = \min_r \frac{1}{2} \| F[m[r] - d \|^2 + \alpha^2 \| Ar \|^2 \]

\[ \nabla J[m] = DF[m](F[m]r - d, r) \]

fact: \( DF[m] \) is badly scaled (unbounded)
Figure 4.3: Model $\delta m + r_k$ (a) and spectrum of model $r_k$ (b) with $k = 30$. 
Figure 4.27: Output of $DF[m][dm, \delta m + r_k]$ with $k = 30$, $m = (2\text{km/s})^2$ and $\delta m$ shown in Figure 4.10b with bandpass filtered source wavelet (a) and source wavelet integrated over time twice (b).
Figure 4.5: Model $\delta m + r_k$ (a) and spectrum of model $r_k$ (b) with $k = 70$. 
Figure 4.28: Output of $DF[m][dm, \delta m + r_k]$ with $k = 70$, $m = (2\text{km/s})^2$ and $\delta m$ shown in Figure 4.10b with bandpass filtered source wavelet (a) and source wavelet integrated over time twice (b).
Chapter 4

Figure 4.30: Quotient of $L_2$ norms of $DF[m][dm, r_k]$ and $F[m]r_k$, when using (a) bandpass filtered source wavelet (Figure 4.1b) and (b) twice integral of bandpass filtered source wavelet (Figure 4.8).
Chapter 4

Order of DF =

Order of $F + 1$

Convergent inner solve not sufficient for convergent computed gradient
Convergent Gradients

Key ingredients:

- parametrix = asymptotic inverse
- robust optimization
Convergent Gradients

Computable parametrices exist -

- subsurface offset extn
- some source extns
Convergent Gradients

Example: subsurface offset acoustic Born (Hou & S. Geophys. 15)

\[ F[v] = \text{modeling op, velo } v \]

\[ F[v]^{\dagger} = \text{asympt inverse} = W_m[v]^{-1}F[v]^T W_d[v] \]
Convergent Gradients

Weight ops $W_m, W_d$ are filters - cheap, no raytracing or PDE solves

Makes $F$ almost unitary in weighted norms
Convergent Gradients

Modeled data $d = F[v]r$
Convergent Gradients

Data residual $F[v]F[v]^{\dagger}d$
Convergent Gradients

means (roughly):

\[ F[v]^\dagger F[v] = I + S[v], \]

\(S[v]\) is \textit{smoothing} (suppresses HF signal) of order -1
Convergent Gradients

Ignoring regularization, $\alpha \to 0$ limit is

$$\min J[v] = \frac{1}{2} \| Ar[v] \|^2 :$$


some algebra (see paper in TRIP16)…
Convergent Gradients

and ignoring second LS problem,

\[ \nabla J[v] = DF[v]^*(d, A^\dagger Ar[v]) \]

trouble: \( r[v] \leftarrow r_{\text{approx}} \)
Convergent Gradients


so

Convergent Gradients

order $+1/-1$:

$$\nabla J[v] \approx DF[v]^* (d, A^\dagger AF[v]^\dagger d)$$

$$+ DF[v]^* (d, A^\dagger AS[v] r_{\text{approx}})$$

First term: Jie’s appinv gradient; second term: correction for inner inversion
Convergent Gradients

can compute $S[v] = I - F[v] \dagger F[v]$!

$[\text{grad error}] \leq [\text{error in } r]$
Convergent Gradients

more huffing and puffing:

[Error in \( \nabla J \) \( \leq \) \([K \times \text{error in normal eqn}]\)]

more trouble: no explicit control of \( K \)
Convergent Gradients

Heinkenschloss-Vicente 01: variant of trust-region qN

step length control: short enough step is near steepest descent so always works
Convergent Gradients

H-V01: converges with inexact grad, provided

\[ \text{grad error} \leq K \times \max(|\text{approx grad}|, \text{step bound}) \]

our case: \[ \text{error in normal eqn} \leq \max(\ldots) \]
Convergent Gradients

Upshot: assure convergence via

- parametrix $\Rightarrow$ control grad error
- couple grad error to step control
Conclusion

Inversion: practical $\Rightarrow$ reliable, efficient

- Yin’s thesis: shot record LSM accel., clarified reliability issues with EFWI
- for separable EFWI: critical requirement is computable parametrix
- couple accuracy and step control
Thanks to...

- our sponsors
- present and former TRIPpers
- many colleagues
- developers of SU, Madagascar
- TACC & RCSG