Data Asssimilation via Piecewise Linearization

Diplomarbeit

Humboldt-Universität zu Berlin
Mathematisch-Naturwissenschaftliche Fakultät II
Institut für Mathematik

eingereicht durch: Ben Lenser
geb am: 27.05.1988 in: Berlin
Betreuer: Prof. PhD A. Griewank
Zweitbetreuer: Prof. Dr. M. Arnold[1]
Berlin, den 12. Juni 2014

[1]Martin-Luther-Universität Halle-Wittenberg

Contents

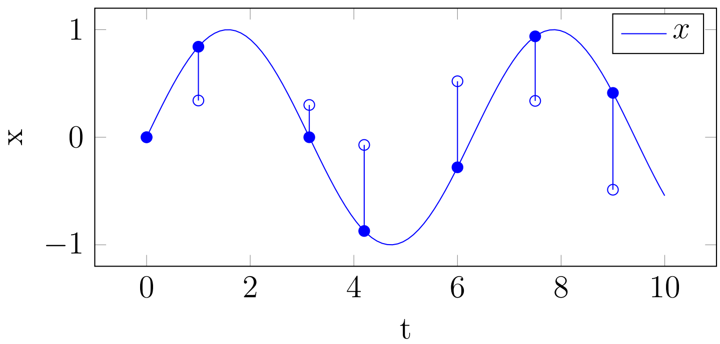
1 Einführung

2 Ordinary Differential Equations

3 Automatische Differentiation und stückweise Linearisierung

4 Datenassimilation

4.1 Problemstellung



Das Ziel der 4-D Datenassimilierung

5 Datenassimilierung mittels stückweiser Linearisierung

5.1 Kinks ermitteln

test

5.2 Verallgemeinerter Gradient

Wenn wir die Stückweise Linearisierung unserer Funktion in Abs Normal Form haben, ist es ein Leichtes, den verallgemeinerten Gradienten zu bestimmen. Dazu betrachten wir die ausmultiplizierte Abs Normal Form

$$z = c + Zx + L|z|$$

Wenn wir uns auf einem Polyeder befinden, sind die Vorzeichen von z eindeutig bestimmt, es ergibt sich also

$$|z| = \Sigma z \quad \text{für} \quad \Sigma_{ii} = \begin{cases} \operatorname{sign}(z_i) & z_i \neq 0 \\ 0 & \text{sonst} \end{cases}, \quad i = 1 \ldots s$$

Umgestellt ergibt sich

$$
\begin{aligned}
& & z &= c + Zx + L\Sigma z \\
&\iff & (I - L\Sigma)z &= c + Zx \\
&\iff & z &= (I - L\Sigma)^{-1}(c + Zx)
\end{aligned}
$$

Das Inverse von $(I - L\Sigma)$ ist wohldefiniert, da Σ eine Diagonalmatrix, L eine strikt untere Dreiecksmatrix und damit $(I - L\Sigma)$ invertierbar ist. Aufgrund der strikt unteren Dreiecksstruktur von L ist die Matrix nilpotent vom Grade $\nu < s$, das bedeutet, dass die Neumannreihe konvergiert und insbesondere endlich ist

$$(I - L\Sigma)^{-1} = I + L\Sigma + (L\Sigma)^2 + \ldots + (L\Sigma)^{\nu-1}$$

Offensichtlich ergibt sich im einfachen Fall mit *switching depth* $\nu = 1$, dass $L = 0$ und damit $(I - L\Sigma)^{-1} = I^{-1} = I$. Im *simply switched*- Fall, $\nu = 2$, ergibt sich immer noch, dass $(I - L\Sigma)^{-1} = I + L\Sigma$. Eingesetzt in Gleichung (??) folgt

$$
\begin{aligned}
y &= b + Jx + Y|z| \\
y &= b + Jx + Y\Sigma z \\
y &= b + Jx + Y\Sigma((I - L\Sigma)^{-1}(c + Zx))z \\
y &= b + Y\Sigma(I - L\Sigma)^{-1}c + \underbrace{(J + Y\Sigma(I - L\Sigma)^{-1}Z)}_{J_\sigma}x
\end{aligned}
$$

wobei gerade J_σ der gesuchten Jacobimatrix von F im Polyeder P_σ entspricht. Griewank et al. haben dieses Ergebnis in [Gri+13, Proposition 2.2] zusammengefasst zu

Theorem 5.1 (Explizite Jacobimatrix Darstellung)
Auf allen offenen P_σ kann y direkt in Abhängigkeit von x ausgedrückt werden als

$$y = b + Y\Sigma(I - L\Sigma)^{-1}c + J_\sigma x \quad \text{mit} \quad J_\sigma = J + Y\Sigma(I - L\Sigma)^{-1}Z \quad (5.1)$$

Hier ist J_σ die Jacobimatrix von F eingeschränkt auf P_σ. Sie reduziert sich zu $J_\sigma = J + Y\Sigma Z$ für einfache Verschachtelungstiefe und zu J für glatte Probleme.

Da die Jacobimatrix an Kinks nicht eindeutig sein muss, wird ein Vektor d mit angegeben, welcher uns die Richtung der Ableitung vorgibt. Diese ist dann wiederum eindeutig bestimmt. Um dies mit unseren eben geführten Betrachtungen zu vereinen berechnen wir also den Gradienten J_σ nicht an der gegebenen Stelle x_0, sondern gehen ein Stück in das durch d und Σ vorgegebene Polyeder und berechnen dort die Ableitung an Stelle $\mathring{x}_0 = x_0 + \frac{1}{2}\tau \cdot d$. Aufgrund der Linearität der Polyeder stimmen die Ableitung an der Stelle x_0 und jene an Stelle $\mathring{x}_0$ überein. Der Algorithmus ergibt sich sofort zu [Bild auf Linker Seite]

Procedure 1 Berechnung des verallgemeinerten Gradienten an Stelle x_0

```
function genJac(x, d)
    τ ← crit_mult(x, d)                                  ▷ 0 < τ < 1, berechne τ an Stelle x in d
    z ← calculate_z(x + 0.5 · τ · d)                     ▷ Berechne Gradient an Mittelpunkt
    Σ ← calc_sigma(z)                                    ▷ Berechne Vorzeichenmatrix Σ
    L_σ ← L · Σ
    L_σ+ ← 0, L_σ* ← I
    repeat                                               ▷ Berechne Neumann Reihe bis Konvergenz
        L_σ+ ← L_σ+ + L_σ*
        L_σ* ← L_σ* · L_σ
    until (L_σ* ≡ 0)                                     ▷ Überprüfe auf Konvergenz
    J_σ ← J + Y · Σ · L_σ+ · Z                           ▷ Berechne Gradienten
    return J_σ
end function
```

5.3 Verallgemeinerte Mittelpunktsregel für das adjungierte Modell

Wie wir in (??) beobachtet haben, müssen wir das Adjungierte Modell rückwärts integrieren. Die Grundidee besteht darin, dass wir auf das Adjungierte Modell (??) der Datenassimilierung ebenfalls wieder die verallgemeinerte Mittelpunktsregel (??) anwenden. Daher, dass wir das Integral in Teilintervalle zerlegen können, berechnen wir die kritischen Multiplikatoren, springen von Kink zu Kink und berechnen auf diesen Teilintervalen den Gradienten unserer rechten Seite F. Sei τ der kritische Multiplikator zum nächsten Kink. x_n und x_a sind die jeweiligen iterierten Werte mit $\mathring{x}$ als deren Mittelpunkt und $\triangle\mathring{x}$ als Differenz. Dann ergibt sich sofort

$$\begin{aligned}
\mathring{x} &= \frac{\hat{x}+\check{x}}{2} \iff \check{x} = 2\mathring{x} - \hat{x} \\
\triangle\mathring{x} &= \mathring{x}_n - \mathring{x}_a \iff \mathring{x}_n = \triangle\mathring{x} + \mathring{x}_a \\
\hat{x} - \check{x} &= 2\hat{x} - 2\mathring{x} \\
\triangle\tau_j &= \tau_j - \tau_{j-1}
\end{aligned}$$

Unsere adjungierte Differentialgleichung mit adjungierter Variable $\bar{x}$ ergibt sich zu

$$\dot{\bar{x}} = x - x_{obs} - \frac{\partial F(x,d)}{\partial x}^T \cdot \bar{x} \tag{5.2}$$

Wenn wir nun die verallgemeinerte Mittelpunktsregel (??) auf 5.2 anwenden, folgt

$$\begin{aligned}
\bar{x}_n - \bar{x}_a &= h \cdot \int_{-0.5}^{0.5} \mathring{x} - \mathring{x}_{Obs} - \frac{\partial F(\mathring{x},d)}{\partial x}^T \cdot \bar{x} \\
&= h \cdot (\mathring{x} - \mathring{x}_{Obs} - \int_{-0.5}^{0.5} \frac{\partial F(\mathring{x},d)}{\partial x}^T \cdot \bar{x} dt)
\end{aligned}$$

Angenommen wir haben $l \in \mathbb{N}$ Kinks zwischen x_n und x_a, wobei $-0.5 = \tau_0 < \tau_1 < \ldots < \tau_l = 0.5$, dann können wir unser Integral aufteilen in

$$\begin{aligned}
\bar{x}_n - \bar{x}_a &= h \cdot (\mathring{x} - \mathring{x}_{Obs} - \sum_{i=1}^{l} \int_{\tau_{i-1}}^{\tau_i} \frac{\partial F(\mathring{x} + \tau_{i-1} d, d)}{\partial x}^T \cdot \bar{x} dt) \\
&= h \cdot (\mathring{x} - \mathring{x}_{Obs} - \sum_{i=1}^{l} \underbrace{\frac{\partial F(\mathring{x} + \tau_{i-1} d, d)}{\partial x}^T}_{A_i^T} \cdot \int_{\tau_{i-1}}^{\tau_i} \bar{x} dt)
\end{aligned}$$

Nun wenden wir die stückweise Linearisierung auf $\bar{x}$ an. Es entsteht

$$
\begin{aligned}
\bar{x}_n - \bar{x}_a &= h \cdot (\mathring{x} - \mathring{x}_{Obs} - \sum_{i=1}^{l} A_i^T \cdot \int_{\tau_{i-1}}^{\tau_i} \mathring{\bar{x}} + t \triangle \bar{x} dt) \\
&= h \cdot (\mathring{x} - \mathring{x}_{Obs} - \sum_{i=1}^{l} A_i^T \cdot \left[t\mathring{\bar{x}} + \frac{t^2}{2} \triangle \bar{x} \right]_{\tau_{i-1}}^{\tau_i}) \\
&= h \cdot (\mathring{x} - \mathring{x}_{Obs} - \sum_{i=1}^{l} A_i^T \cdot \Big((\tau_i - \tau_{i-1}) \cdot \mathring{\bar{x}} + \frac{1}{2} \triangle \bar{x} \cdot (\tau_i^2 - \tau_{i-1}^2) \Big)) \\
&= h \cdot (\mathring{x} - \mathring{x}_{Obs} - \sum_{i=1}^{l} A_i^T \cdot \Big(\triangle\tau_i \cdot \mathring{\bar{x}} + \triangle\tau_i \frac{1}{2} \cdot (\tau_i + \tau_{i-1}) \cdot \triangle\bar{x} \Big)) \\
&= h \cdot (\mathring{x} - \mathring{x}_{Obs} - \sum_{i=1}^{l} A_i^T \cdot \big(\triangle\tau_i \cdot \mathring{\bar{x}} + \triangle\tau_i \mathring{\tau}_i \triangle \bar{x} \big))
\end{aligned}
$$

Mit $\mathring{\bar{x}} = \frac{1}{2}(\bar{x}_n + \bar{x}_a)$ und $\triangle\bar{x} = \bar{x}_n - \bar{x}_a$ ergibt sich

$$
\begin{aligned}
\bar{x}_n - \bar{x}_a &= h \cdot (\mathring{x} - \mathring{x}_{Obs} - \sum_{i=1}^{l} A_i^T \cdot \Big[\triangle\tau_i \cdot \frac{\bar{x}_n + \bar{x}_a}{2} + \triangle\tau_i \mathring{\tau}_i (\bar{x}_n - \bar{x}_a) \Big]) \\
&= h \cdot (\mathring{x} - \mathring{x}_{Obs} - \sum_{i=1}^{l} A_i^T \cdot \Big[\Big(\frac{1}{2} \triangle \tau_i + \triangle\tau_i \mathring{\tau}_i \Big) \cdot \bar{x}_n + \Big(\frac{1}{2} \triangle \tau_i - \triangle\tau_i \mathring{\tau}_i \Big) \cdot \bar{x}_a \Big]) \\
&= h \cdot (\mathring{x} - \mathring{x}_{Obs} - \Big[\sum_{i=1}^{l} A_i^T \Big(\frac{1}{2} \triangle \tau_i + \triangle\tau_i \mathring{\tau}_i \Big) \cdot \bar{x}_n + \sum_{i=1}^{l} A_i^T \Big(\frac{1}{2} \triangle \tau_i - \triangle\tau_i \mathring{\tau}_i \Big) \cdot \bar{x}_a \Big])
\end{aligned}
$$

Durch Umsortierung erhalten wir

$$
\begin{aligned}
(I + h \sum_{i=1}^{l} A_i^T \Big(\frac{1}{2} \triangle \tau_i + \triangle\tau_i \mathring{\tau}_i \Big)) \bar{x}_n =& \ (I - h \sum_{i=1}^{l} A_i^T \Big(\frac{1}{2} \triangle \tau_i - \triangle\tau_i \mathring{\tau}_i \Big) \bar{x}_a \\
& + h \cdot (\mathring{x} - \mathring{x}_{Obs}) \\
\iff \quad \Big(I + \frac{h}{2} \bar{A}^T + h\hat{A}^T \Big) \bar{x}_n =& \Big(I - \frac{h}{2} \bar{A}^T + h\hat{A}^T \Big) \bar{x}_a + h \cdot (\mathring{x} - \mathring{x}_{Obs})
\end{aligned}
$$

mit $\bar{A}^T = \sum_{i=1}^{l} A_i^T \triangle \tau_i$ und $\hat{A}^T = \sum_{i=1}^{l} A_i^T \triangle \tau_i \mathring{\tau}_i$. Als Algorithmus ergibt sich folglich

Procedure 2 `PlanC::calc_kink_partials`

function CALC__KINK__PARTIALS$(\check{x}, \hat{x}, d)$
    $\hat{\tau} \leftarrow 0,\ x_{kink} \leftarrow \check{x}$
    $\bar{A} \leftarrow 0,\ \hat{A} \leftarrow 0$
    $d \leftarrow \frac{d}{\|d\|} \cdot \|\hat{x} - \check{x}\|$ ▷ Normalize direction
    repeat
        $\check{\tau} \leftarrow \hat{\tau},\ x_{kink} \leftarrow x_{kink} + \check{\tau} d$
        $\hat{\tau} \leftarrow$ CRITMULT(x_{kink}, d) ▷ Berechne kritischen Multiplikator bis zum nächsten Kink
        if $\hat{\tau} > 1$ **then** $\hat{\tau} \leftarrow 1$
        $\mathring{x} \leftarrow x_{kink} + 0.5 \cdot \hat{\tau} d$ ▷ Berechne Mittelpunkt zwischen den Kinks
        $\frac{\partial F(\mathring{x})}{\partial x} \leftarrow \mathrm{gen\_jac}(\mathring{x}, d)$ ▷ Berechne ∂F aus der Abs-Normalform
        $\bar{A} \leftarrow \bar{A} + \frac{\partial F(\mathring{x})}{\partial x} \cdot (\hat{\tau} - \check{\tau})$
        $\hat{A} \leftarrow \hat{A} + \frac{\partial F(\mathring{x})}{\partial x} \cdot \frac{1}{2}(\check{\tau} + \hat{\tau}) - 0.5$ ▷ Verschiebe τ um -0.5
    until $\hat{\tau} \geq 1$
    return $[\bar{A}, \hat{A}]$;
end function

Procedure 3 `PlanC::jac_data_assimilation`

Input: $x_0, t_0, T, h, x_{Obs}, TOL$

$N = \left\lceil \frac{t_0 - T}{h} \right\rceil$

$\hat{\mathring{x}} \leftarrow 0$ — Setze Anfangswert

$x \leftarrow \text{SOLVEODE}(x_0, t_0, T, h, TOL)$; — Löse ODE in Vorwärtsrichtung

for $k \leftarrow$ N-1 to 1 **do** — Zeitschritt rückwärts

  $\mathring{x} \leftarrow 0.5(x_k + x_{k-1})$ — Berechne Mittelpunkt

  $\square_{\mathring{x}} F \leftarrow \text{UPDATE}$ — Berechne neue Linearisierung am Mittelpunkt $\mathring{x}$

  $d \leftarrow x_{k-1} - x_k$ — Berechne neue Richtung

  $[\bar{A}, \hat{A}] \leftarrow \text{CALC\_\_KINK\_\_PARTIALS}(x_k, x_{k-1}, d)$ — Berechne ∂F zwischen jedem Kink

  $\check{\mathring{x}} \leftarrow \text{SOLVE}((I - \frac{h}{2}\bar{A}^T + h\hat{A}^T)\check{\mathring{x}} = (I + \frac{h}{2}\bar{A}^T + h\hat{A}^T)\hat{\mathring{x}} - h(\mathring{x} - \mathring{x}_{Obs}))$

end for

return $-\check{\mathring{x}}$

Bemerkt sei, dass sich die Formel für $l = 1$, also kein Kink zwischen x_n und x_a, zur bekannten impliziten Mittelpunktsregel vereinfacht. Für diesen Fall gilt $\triangle\tau = 1, \mathring{\tau} = 0$ und damit

$$
\begin{aligned}
& (I + h\sum_{i=1}^{1} A_i^T \left(\frac{1}{2} \triangle \tau_i + \triangle\tau_i\mathring{\tau}_i\right))\bar{x}_n = (I - h\sum_{i=1}^{1} A_i^T \left(\frac{1}{2} \triangle \tau_i - \triangle\tau_i\mathring{\tau}_i\right)\bar{x}_a \\
& \qquad + h \cdot (\mathring{x} - \mathring{x}_{Obs}) \\
\iff \quad & (I + hA_1^T \left(\frac{1}{2} \triangle \tau_1 + \triangle\tau_1\mathring{\tau}_1\right))\bar{x}_n = (I - hA_1^T \left(\frac{1}{2} \triangle \tau_1 - \triangle\tau_1\mathring{\tau}_1\right)\bar{x}_a \\
& \qquad + h \cdot (\mathring{x} - \mathring{x}_{Obs}) \\
\iff \quad & (I + hA_1^T \left(\frac{1}{2} \cdot 1 + 0\right))\bar{x}_n = (I - hA_1^T \left(\frac{1}{2} \cdot 1 - 0\right)\bar{x}_a + h \cdot (\mathring{x} - \mathring{x}_{Obs}) \\
\iff \quad & \left(I + \frac{h}{2}\frac{\partial F(\mathring{x}, d)}{\partial x}^T\right)\bar{x}_n = \left(I - \frac{h}{2}\frac{\partial F(\mathring{x}, d)}{\partial x}^T\right)\bar{x}_a + h \cdot (\mathring{x} - \mathring{x}_{Obs}) \\
\iff \quad & \bar{x}_n - \bar{x}_a = h \cdot \left(\mathring{x} - \mathring{x}_{Obs} - \frac{\partial F(\mathring{x}, d)}{\partial x}^T \frac{\bar{x}_n + \bar{x}_a}{2}\right)
\end{aligned}
$$

vgl. [Satz wo normale implizite midpointrule aufgeführt wird]

6 Lösen von ODEs mittels stückweiser Linearisierung

7 Experiments

7.1 Shallow Water Equation

8 Implementation

Bibliography

[Gri+13] Andreas Griewank et al. *Solving piecewise linear equations in abs-normal form.* 2013.