Optimale Steuerung

Sequentielle Quadratische Programmierung

Kevin Sieg

Fachbereich für Mathematik und Statistik
Universität Konstanz

14. Juli 2010

Aufgabenstellung

Löse folgendes Minimierungsproblem:

$$\min J(y,u) := \frac{1}{2}\int_\Omega |y(x) - y_d(x)|^2 dx + \frac{\gamma}{2}\|u - u_d\|_U^2 \tag{1}$$

unter den Nebenbedingungen (mit $U = \mathbb{R}^2$)

$$-\Delta y(x) + a(x)y(x) + y^3(x) = \sum_{i=1}^{\mathrm{L}} u_i b_i(x) \quad \forall x \in \Omega \; u \in \mathbb{R}^L \tag{2}$$

$$y(x) = 0 \quad \forall x \in \delta\Omega \tag{3}$$

Die zusammengesetzte Trapezregel I

Näherungsweise Berechnung eines Integrals $[a, b]$ durch Zerlegung in M Teilintervalle:
Teile Intervall in $a = x_0 < x_1 < \cdots < x_{M-1} < x_M = b$ ein:

$$\int_a^b f(t)dt = \sum_{i=1}^{N} \int_{x_{i-1}}^{x_i} f(t)dt.$$

Teilintervalle gleiche Größe: $x_i := a + i \, \frac{b-a}{M} \quad (i = 0, \ldots, M)$
Jeweiliges Teilintervall $[x_{i-1}, x_i]$: einfache Trapezregel:

$$\int_{x_{i-1}}^{x_i} f(t)dt = \frac{b-a}{M} \big(f(x_{i-1}) + f(x_i) \big) + R_i[f]$$

Die zusammengesetzte Trapezregel II

Somit ergibt sich für das Gesamtintervall folgende Quadraturformel:

$$\int_a^b f(t)dt = \frac{b-a}{M}\left(\frac{1}{2}f(a) + \sum_{i=1}^{M-1} f(x_i) + \frac{1}{2}f(b)\right) + R[f].$$

Anwendung auf das Zielfunktional

Die Zielfunktion

$$J(y,u) = \frac{1}{2}\int_\Omega |y(x) - y_d(x)|^2 dx + \frac{\gamma}{2}\|u - u_d\|_U^2 \qquad (4)$$

lässt sich nun wie folgt diskretisieren und umschreiben:
Da das Gebiet $\Omega = (0,1)^2$ ist, gilt $h = 1/M$. Sei nun

$$y = (y^1, \dots, y^{M-1}) \qquad y^j = (y_{1,j}, \dots, y_{M-1,j})$$
$$y_d = (y_d^1, \dots, y_d^{M-1}) \qquad y_d^j = (y_{d_{1,j}}, \dots, y_{d_{M-1,j}})$$

wobei $y(ih, jh) = y_{ij}$ ist. Das Verfahren lässt sich mit

$$Q = h^2 \cdot \mathrm{I}_{(M-1)^2} (\text{Rand ist null!})$$

darstellen:

$$\widetilde{J}(y,u) = \frac{1}{2}(y - y_d)^T Q (y - y_d) + \frac{\gamma}{2}(u - u_d)^T (u - u_d). \qquad (5)$$

Das finite Differenzenverfahren

Ersetze Ableitungen durch symmetrische Differenzen:
Approximation der ersten Ableitung:

$$y'(t) = \frac{y(t+h) - y(t-h)}{2h}$$

Approximation der zweiten Ableitung:

$$y''(t) = \frac{1}{h^2}\big(y(t-h) - 2y(t) + y(t+h)\big) + \mathcal{O}(h^2).$$

Der Laplaceoperator lässt sich dann wie folgt approximieren:

$$-(\Delta y)_{ij} \approx \frac{1}{h^2}(-y_{i-1,j} - y_{i+1,j} + 4y_{ij} - y_{i,j-1} - y_{i,j+1}). \quad (6)$$

Dabei sei $y_{ij} = y(ih, jh)$.

Haben also Matrizen :

$$A_{\text{laplace}} := \begin{pmatrix} B & -C & 0 & \dots & 0 & 0 & 0 \\ -C & B & -C & \dots & 0 & 0 & 0 \\ 0 & -C & B & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & B & -C & 0 \\ 0 & 0 & 0 & \dots & -C & B & -C \\ 0 & 0 & 0 & \dots & 0 & -C & B \end{pmatrix} \in \mathbb{R}^{(}M-1)^2\times(M-$$

mit

$$B = \frac{1}{h^2}\begin{pmatrix} 4 & -1 & 0 & \dots & 0 & 0 & 0 \\ -1 & 4 & -1 & \dots & 0 & 0 & 0 \\ 0 & -1 & 4 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 4 & -1 & 0 \\ 0 & 0 & 0 & \dots & -1 & 4 & -1 \\ 0 & 0 & 0 & \dots & 0 & -1 & 4 \end{pmatrix} \quad \text{und} \quad C = \frac{1}{h^2}\mathrm{I}_{\mathbb{R}^{M-1\times M-}}$$

Diskretisierung der Nebenbedingungen

$$\min_{(y,u)} J(y,u) = \frac{1}{2}(y-y_d)^T Q(y-y_d) + \frac{\gamma}{2}(u-u_d)^T(u-u_d) \tag{7a}$$

unter der Nebenbedingung

$$e(y,u) := Ay + H(y) - Bu = 0 \tag{7b}$$

Dabei ist $H : \mathbb{R}^{(M-1)^2} \to \mathbb{R}^{(M-1)^2}$; $y \mapsto y^3$ und $B = (b_1(x), \ldots, b_L(x))$ eine Gebietsteuerung auf den Teilgebieten Ω_i mit $b_i(x) = \chi_{\Omega_i}$ ausgewertet an den Stützstellen und $A = A_{\text{laplace}} + a(x)$ mit $a(x)$ aus der Nebenbedingung ebenfalls ausgewertet an den Stützstellen (Diagonalmatrix).

Lagrangefunktion des Systems

Lagrangefunktion

$$\mathcal{L}(y,u,p) = \frac{1}{2}(y-y_d)^T Q(y-y_d) + \frac{\gamma}{2}(u-u_d)^T(u-u_d)$$
$$+p^T(Ay + H(y) - Bu),$$

oder mit $z := (y,u) \in \mathbb{R}^{(M-1)^2} \times \mathbb{R}^L$ und $p \in \mathbb{R}^{(M-1)^2}$ kurz

$$\mathcal{L}(z,p) = J(z) + p^T e(z).$$

wobei $J(z) = (y-y_d)^T Q(y-y_d) + \frac{\gamma}{2}(u-u_d)^T(u-u_d)$.
Schreibe im Folgenden $n := (M-1)^2$ und $m := L$.

KKT-Bedingungen

Die Karush-Kuhn-Tucker Bedingungen $\nabla\mathcal{L}(z,p)=0$ führen zu folgendem Gleichungssystem:

$$\begin{aligned} Q(y-y_d)+(A+H'(y))^T p &= 0 \\ \gamma(u-u_d)-B^T &= 0 \\ Ay+H(y)-Bu &= 0 \end{aligned} \tag{8}$$

beziehungsweise kompakt mit der Notation $e'(z)=(e_y(z),e_u(z))\in\mathbb{R}^{n\times(n+m)}$

$$F(z,p)=\nabla\mathcal{L}(z,p)=\begin{pmatrix}\nabla J(z)+e'(z)^T p \\ e(z)\end{pmatrix}\overset{!}{=}0. \tag{9}$$

wobei gilt

$$e_y(z)=A+H'(y) \quad \text{und} \quad e_u(z)=-B.$$

Newton-Verfahren

Newton-Verfahren

Iterationsvorschrift des Newton-Verfahrens

$$x^{k+1} := x^k - ((F'(x^k))^{-1} F(x^k)$$

Bei der praktischen Durchführung löst man das lineare System

$F'(x^k)d^k = -F(x^k)$ und setzt $x^{k+1} = x^k + d^k, \quad (k = 0, 1, 2, \dots).$

Anwendung des Newton Verfahrens auf das KKT-System I

Wende nun das Newton-Verfahren an, um das KKT-System (9) zu lösen. Die Jacobi-Matrix von F ist

$$F'(z,p) = \nabla^2 \mathcal{L}(z,p) = \begin{pmatrix} \mathcal{L}_{zz}(z,p) & e'(z)^T \\ e'(z) & 0 \end{pmatrix},$$

wobei

$$\mathcal{L}_{zz}(z,p) = \begin{pmatrix} \mathcal{L}_{yy}(z,p) & 0 \\ 0 & \mathcal{L}_{uu}(z,p) \end{pmatrix} = \begin{pmatrix} Q + H''(y)^T p & 0 \\ 0 & \gamma \, \mathrm{I}_{R^m} \end{pmatrix}.$$

Es gilt also in jedem Newton-Schritt folgendes System zu lösen:

$$\nabla^2 \mathcal{L}(z^k, p^k) \begin{pmatrix} \Delta z^k \\ \Delta p^k \end{pmatrix} = -\nabla \mathcal{L}(z^k, p^k) \tag{10a}$$

mit dem Update

$$\begin{pmatrix} z^{k+1} \\ p^{k+1} \end{pmatrix} = \begin{pmatrix} z^k \\ p^k \end{pmatrix} + \begin{pmatrix} \Delta z^k \\ \Delta p^k \end{pmatrix} \tag{10b}$$

Umformulierung des KKT-Systems

Multipliziere den Schritt des Newton-Systems aus:

$$\begin{aligned} \mathcal{L}_{zz}(z^k,p^k)\Delta z^k + \underbrace{e'(z^k)^T\Delta p^k}_{=e'(z^k)^T(p^{k+1}-p^k)} &= -\nabla J(z^k) - e'(z^k)^T p^k \\ e'(z^k)\Delta z^k &= -e(z^k) \end{aligned}$$

wird duch Kürzung zu

$$\begin{aligned} \mathcal{L}_{zz}(z^k,p^k)\Delta z^k + e'(z^k)^T p^{k+1} &= -\Delta J(z^k) \\ e'(z^k)\Delta z^k &= -e(z^k) \end{aligned}$$

Setze nun $\mu^k = p^{k+1}$:

$$\begin{pmatrix} \mathcal{L}_{zz}(z^k,p^k) & e'(z^k)^T \\ e'(z^k) & 0 \end{pmatrix} \begin{pmatrix} \Delta z^k \\ \mu^k \end{pmatrix} = -\begin{pmatrix} J'(z^k) \\ e(z^k) \end{pmatrix}. \tag{11}$$

Definiere eine neue Lagrange-Funktion (diese hat quadratische From):

$$\widetilde{\mathcal{L}}(z^k, \mu^k) := \frac{1}{2}(\Delta z^k)^T \mathcal{L}_{zz}(z^k, p^k)\Delta z^k + \nabla J(z^k)\Delta z^k +$$

$$\mu^k(e'(z^k)\Delta z^k + e(z^k)),$$

$\rightarrow$ habe in (11) gerade die notwendige Bedingung erster Ordnung

$$\nabla \widetilde{\mathcal{L}}(z^k, \mu^k) = 0 \tag{12a}$$

folgenden quadratischen Teilproblems darstellt:

$$\min_{\Delta z^k} \frac{1}{2}(\Delta z^k)^T \mathcal{L}_{zz}(z^k, p^k)\Delta z^k + \nabla J(z^k)\Delta z^k \tag{12b}$$

unter der Nebenbedingung

$$e'(z^k)\Delta z^k + e(z^k) = 0, \tag{12c}$$

- In jedem Schritt werden quadratische Subprobleme gelöst
- Annahmen für die Lösbarkeit
  - Die Jacobi-Matrix $e'(z^k)$ hat vollen Zeilenrang
  - Die Matrix $\mathcal{L}_{zz}(z^k, p^k)$ ist positiv definit auf dem Kern von $e'(z^k)$,das heißt

$$d^T \mathcal{L}_{zz}(z^k, p^k) d > 0$$

für alle $d \neq 0$ mit $e'(z^k)d = 0$.

SQP-Algorithmus

Lokales SQP-Verfahren

1. **Wähle geeignete Startwerte**

$$z^0 \in \mathbb{R}^{n+m},\ p^0 \in \mathbb{R}^n,\ k^0 = 0 \textbf{ sowie } k_{\max} \in \mathbb{N}.$$

2. **while** $(k \leq k_{\mathbf{max}})$
   - **berechne**

$$\begin{aligned} J^k &= J(z^k), & \nabla J^k &= \nabla J(z^k), & \mathcal{L}_{zz}^k &= \mathcal{L}_{zz}(z^k, p^k), \\ e^k &= e(z^k) & \nabla e^k &= \nabla e(z^k); \end{aligned}$$

   - **löse (12) für** $(\Delta z^k, \mu^k)$
   - **setze** $z^{k+1} = z^k + \Delta z^k$ **sowie** $p^{k+1} = \mu^k$.
   - **überprüfe, ob weitere Abbruchkriterien erfüllt sind.**

end

Lösung mit der Nullraum-Methode

KKT-System

Das zu lösende KKT-System aus (11) sieht folgendermaßen aus:

$$\begin{pmatrix} \mathcal{L}_{zz}(z^k, p^k) & e'(z^k)^T \\ e'(z^k) & 0 \end{pmatrix} \begin{pmatrix} \Delta z \\ \Delta p \end{pmatrix} = - \begin{pmatrix} \nabla J(z^k) \\ e(z^k) \end{pmatrix},$$

wobei

$$e'(z^k) = (e_y(z^k), e_u(z^k)) \in \mathbb{R}^{n \times (n+m)}$$
$$\text{und} \quad \Delta z = (\Delta y, \Delta u) \in \mathbb{R}^n \times \mathbb{R}^m.$$

Lösung mit der Nullraum-Methode

Suche eine Matrix $Z \in kern(e'(z))$ mit $Z^T Q Z \geq 0$ und Y so, dass $[Y|Z]$ regulär ist. Zerlege $\Delta z = T \Delta z_T + Z \Delta z_z$:
Definiere

$$T(z) := \begin{pmatrix} e_y(z)^{-1} \\ 0 \end{pmatrix} \in \mathbb{R}^{(m+n)\times n} \tag{13}$$

und

$$Z(z) := \begin{pmatrix} e_y(z)^{-1} e_u(z) \\ I_{\mathbb{R}^m} \end{pmatrix} \tag{14}$$

Es gelten folgende Eigenschaften:

$$e'(z)Z(z) = 0 \in \mathbb{R}^{n+m} \quad \text{also ist } Bild(Z) \in Kern(e'(z))$$
$$e'(z)T(z) = I_{\mathbb{R}^n}$$

Zerlegung von Δz:

$$\Delta z = Z(z^k)\Delta u - T(z^k)e(z^k),$$

löst zweite Gleichung im System.

Einsetzen in die erste Gleichung des Systems:

$$\mathcal{L}_{zz}(z^k, p^k)\Delta z + e'(z^k)^T \Delta p = -\nabla J(z^k)$$

führt mit Multiplikation mit $Z(z^k)^T$ von links zu

$$Z(z^k)^T \mathcal{L}_{zz}(z^k, p^k) Z(z^k)\Delta u + (e'(z^k)Z(z^k))^T \Delta p = \\ Z(z^k)^T \Big(-\nabla J(z^k) + \mathcal{L}_{zz}(z^k, p^k) T(z^k) e(z^k) \Big)$$

Löse KKT-System mit der Nullraum-Methode:

(i) Berechne Δu

$$H^k \Delta u = Z(z^k)^T \Big(-\nabla J(z^k) + \mathcal{L}_{zz}(z^k, p^k) T(z^k) e(z^k) \Big)$$

mit

$$H^k = Z(z^k)^T \mathcal{L}_{zz}(z^k, p^k) Z(z^k)$$

(ii) Setze für Δy

$$\Delta x = Z(x^k)\Delta u - T(z^k)e(z^k)$$

(iii) berechne Δp

$$e'(z^k)^T \Delta p = -\nabla J(z^k) - \mathcal{L}_{zz}(z^k, p^k)\Delta z$$

CG-Verfahren

Es gilt $Ax = b$ zu lösen mit A symmetrisch und positiv definit
Gegeben: tol, x^0

Conjugated-gradient Verfahren

initialisiere $r^0 = b - Ax^0, k = 0$ und $p^0 = r^0$

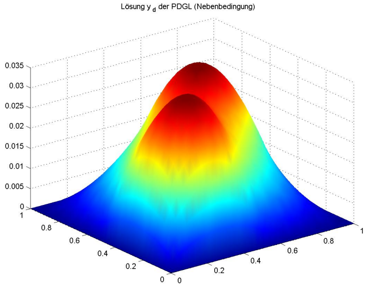
while $(\|r\| \leq tol)$

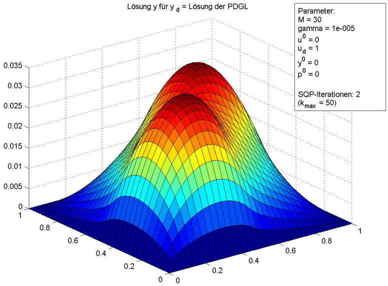
$$\begin{aligned}
a_k &= \frac{(p^k)^T r^k}{(p^k)^T A p^k} \\
x^{k+1} &= x^k + a_k p^k \\
r^{k+1} &= r^k - a_k A p^k \\
\beta_k &= \frac{(Ap^k)^T r^{k+1}}{(Ap^k)^T p^k} \\
p^{k+1} &= r^{k+1} - \beta_k p^k
\end{aligned}$$

end

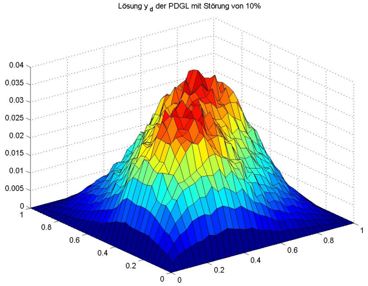
Anhang: Bilder

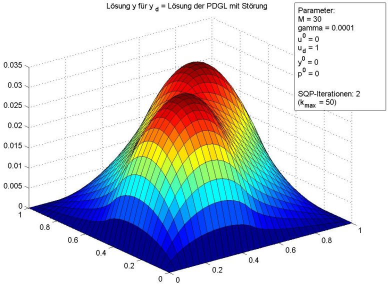
Hier noch ein paar Bilder für verschiedende Funktionen y_d und verschiedene Startwerte. Bei dem y_d mit Störung sei erwähnt, dass die Anzahl der Iterationen natürlich von der zufälligen Störung abhängt, die jedes mal anders sein kann.



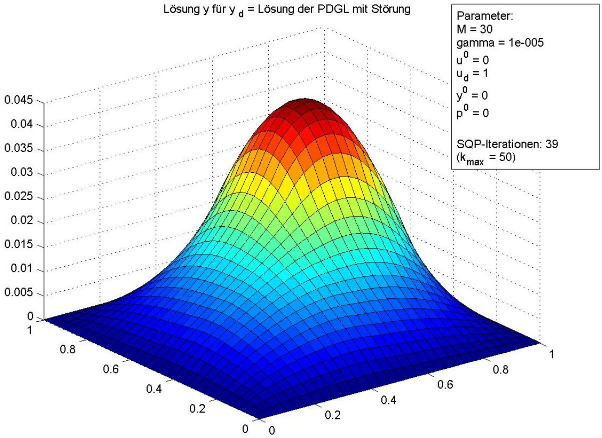


Hier ein sieht man ein y_d mit einer Störung von 10 Prozent, welches ganz gut „geglättet“ wird für entsprechende γ.

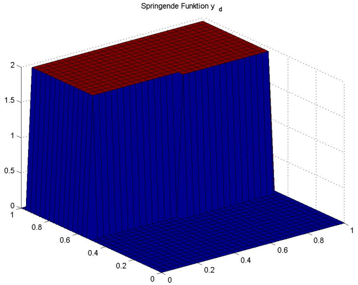


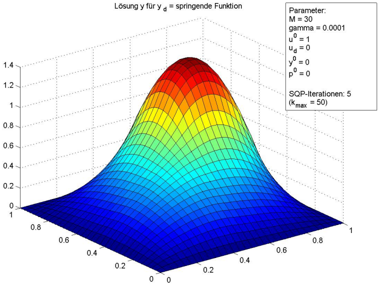


Hier sieht man schön das unterschiedliche Verhalten der Lösung mit unterschiedlichen γ für das y_d mit Störung, besonders im Hinblick auf die Anzahl der Iterationen und die Lösung.



Wählt man y_d als eine Art Treppenfunktion,so sieht man schön, wie der Laplace-Operator der PDGL das ganze glättet und „rund“ macht. Ebenfalls erkennt man gut den Einfluss der Steuerung. Jedoch ist hier die Größenordung der Lösung stark abhängig von γ.





Danke für die Aufmerksamkeit!