

<i>Inhaltsverzeichnis</i>	1
---------------------------	---

Inhaltsverzeichnis

Innere-Punkte-Verfahren	3
--------------------------------	----------

1 Theoretische Grundlagen	3
----------------------------------	----------

1.1 Die KKT-Bedingungen	3
-----------------------------------	---

1.2 Der zentrale Pfad	4
---------------------------------	---

1.3 Existenz des zentralen Pfades	5
---------------------------------------------	---

1.4 Wiederholung: Das Newton-Verfahren	7
--------------------------------------------------	---

2 Innere-Punkte-Verfahren	7
----------------------------------	----------

2.1 Ein zulässiges Verfahren	8
----------------------------------------	---

3 Quellcode	9
--------------------	----------

Innere-Punkte-Verfahren

Gunnar Kedenburg

1 Theoretische Grundlagen

1.1 Die KKT-Bedingungen

Definition 1.1 (Lagrange Funktion) *Die Abbildung*

$$L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \mapsto \mathbb{R},$$

$$L(x, \alpha, \beta) := f(x) + \alpha^T g(x) + \beta^T h(x)$$

heißt Lagrange-Funktion des restringierten Optimierungsproblems

$$\min f(x) \text{ u.d.N. } g(x) \leq 0, h(x) = 0$$

mit stetig differenzierbarer Zielfunktion $f : \mathbb{R}^n \mapsto \mathbb{R}$, und stetig differenzierbaren Funktionen $g : \mathbb{R}^n \mapsto \mathbb{R}^m$ und $h : \mathbb{R}^n \mapsto \mathbb{R}^p$, die die Nebenbedingungen repräsentieren.

Die Lagrange-Funktion macht restringierte Probleme zu unrestringierten Problemen durch Penalty-Terme.

Definition 1.2 (Karush-Kuhn-Tucker-Bedingungen (KKT-Bedingungen))

$$\begin{aligned} \nabla_x L(x, \alpha, \beta) = \nabla f_x + \alpha^T \nabla g(x) + \beta^T \nabla h(x) &= 0 \\ g(x) &\leq 0, h(x) = 0 \\ \alpha &\geq 0, \alpha^T g(x) = 0 \end{aligned}$$

Jeder Vektor (x^*, α^*, β^*) , der den KKT-Bedingungen genügt, heißt KKT-Punkt des Optimierungsproblems. α^* und β^* werden Lagrange-Multiplikatoren genannt.

Bemerkung 1.3 *Für konvexe Optimierungsprobleme mit linearen Restriktionen ist $x^* \in \mathbb{R}^n$ genau dann ein Minimum, wenn es Lagrange-Multiplikatoren $\alpha^* \in \mathbb{R}^m$ und $\beta^* \in \mathbb{R}^p$ gibt, so dass (x^*, α^*, β^*) ein KKT-Punkt ist.*

1.2 Der zentrale Pfad

Betrachte das primale lineare Programm

$$\min c^T x \text{ u.d.N. } Ax = b, x \geq 0$$

und das dazugehörige duale lineare Programm

$$\max b^T \lambda \text{ u.d.N. } A^T \lambda \leq c$$

Die innere Punkte Verfahren sind primal-duale Verfahren. Sie arbeiten zugleich am primalen und dualen Programm.

Definition 1.4 (Optimalitätsbedingungen)

$$\begin{aligned} A^T \lambda + s &= c, \\ Ax &= b \\ x_i s_i &= 0, \quad i = 1, \dots, n \\ x, s &\geq 0 \end{aligned}$$

Satz 1.5 *Erfüllt ein Punkt (x, λ, s) die oben genannten Optimalitätsbedingungen, so gibt es Lagrange-Multiplikatoren α und β , so dass (x, α, β) den KKT-Bedingungen genügt.*

Beweis: Stelle nochmal die KKT-Bedingungen auf:

$$\begin{aligned} \nabla_x L(x, \alpha, \beta) &= \nabla f_x + \alpha^T \nabla g(x) + \beta^T \nabla h(x) = 0 \\ g(x) &\leq 0, \quad h(x) = 0 \\ \alpha &\geq 0, \quad \alpha^T g(x) = 0 \end{aligned}$$

Hierbei ist für unser Optimierungsproblem

$$\begin{aligned} f(x) &= c^T x \\ h(x) &= Ax - b \\ g(x) &= -x \end{aligned}$$

Damit folgt $g(x) \leq 0$ und $h(x) = 0$ direkt aus den Optimalitätsbedingungen. Ausserdem gilt

$$\begin{aligned} \nabla_x L(x, \alpha, \beta) &= \nabla f(x) + \alpha^T \nabla g(x) + \beta^T \nabla h(x) \\ &= c^T - \alpha^T I + \beta^T A \\ &= c^T - s^T - \lambda^T A = 0 \end{aligned}$$

mit $\alpha := s$ und $\beta := -\lambda$. Der letzte Umformungsschritt folgt mit $c = s + A^T \lambda$.

Es gilt $\alpha \geq 0$ wegen $s \geq 0$. Schliesslich folgt $\alpha^T g(x) = s^T(-x) = 0$ mit $x_i s_i = 0$ für alle $i = 1 \dots n$.

□

Definition 1.6 (Zentrale-Pfad-Bedingungen)

$$\begin{aligned} A^T \lambda + s &= c, \\ Ax &= b \\ x_i s_i &= \tau, \quad i = 1, \dots, n, \quad \tau > 0 \\ x, s &> 0 \end{aligned}$$

Die Zentrale-Pfad-Bedingungen sind also gestörte Optimalitätsbedingungen:

Definition 1.7 (Zentraler Pfad) Ist $(x_\tau, \lambda_\tau, s_\tau)$ die Lösung der zentralen Pfad-Bedingungen für ein $\tau > 0$, so nennt man die Abbildung $\tau \mapsto (x_\tau, \lambda_\tau, s_\tau)$ den zentralen Pfad. Dass diese Lösung existiert und eindeutig ist, wird jetzt gezeigt.

1.3 Existenz des zentralen Pfades

Definiere sogenannte Barriere-Probleme zum primalen und dualen Problem:

$$\begin{aligned} \min c^T x - \tau \sum_{i=1}^n \log(x_i) \quad \text{u.d.N.} \quad Ax = b, \quad x > 0 \\ \max b^T \lambda + \tau \sum_{i=1}^n \log(s_i) \quad \text{u.d.N.} \quad A^T \lambda + s = c, \quad s > 0 \end{aligned}$$

Satz 1.8 Sei $\tau > 0$ gegeben. Dann sind die folgenden Aussagen äquivalent:

- (a) Das primale Barriere-Problem besitzt eine Lösung x_τ .
- (b) Das duale Barriere-Problem besitzt eine Lösung (λ_τ, s_τ) .
- (c) Die zentralen Pfad-Bedingungen besitzen eine Lösung (x_τ, λ_τ) .

Beweis: (a) \Leftrightarrow (c):

Die Zielfunktion des primalen Barriere-Problems ist offenbar konvex, und die Ungleichung $x > 0$ aufgrund des Logarithmus immer erfüllt.

Also ist x_τ genau dann eine Lösung des primalen Barriere-Problems, wenn ein λ_τ existiert, so dass die KKT-Bedingungen in $(x, \lambda) = (x_\tau, \lambda_\tau)$ erfüllt sind:

$$\begin{aligned} c - \tau X^{-1} e - A^T \lambda &= 0 \\ Ax &= b \\ x &> 0, \end{aligned}$$

Hierbei ist $e := (1, \dots, 1)^T$ und $X := \text{diag}(x_1, \dots, x_n)$. Mit $s_\tau := \tau X_\tau^{-1} e > 0$ folgt die Aussage.

(b) \Leftrightarrow (c) ist ähnlich. □

Definition 1.9 (primal-dual strikt zulässige Menge)

$$F^0 := \{(x, \lambda, s) \mid Ax = b, A^T \lambda + s = c, x, s > 0\}$$

Satz 1.10 *Ist die strikt zulässige Menge F^0 nichtleer, so besitzt das primale Barriere-Problem für jedes $\tau > 0$ eine Lösung x_τ .*

Beweis: Sei $\tau > 0$ fest gewählt, sowie $(\hat{x}, \hat{\lambda}, \hat{s}) \in F^0$ gegeben.

Zeige: $L_\tau := \{x \in \mathbb{R}^n : Ax = b, x \geq 0, B_\tau(x) \leq B_\tau(\hat{x})\}$ ist kompakt. Abgeschlossenheit ist klar, zeige Beschränktheit:

Sei $x \in L_\tau$.

$$\begin{aligned} B_\tau(x) &= c^T x - \tau \sum_{i=1}^n \log(x_i) \\ &= c^T x - \hat{\lambda}^T (Ax - b) - \tau \sum_{i=1}^n \log(x_i) \\ &= c^T x - x^T A^T \hat{\lambda} + b^T \hat{\lambda} - \tau \sum_{i=1}^n \log(x_i) \\ &= c^T x - x^T (c - \hat{s}) + b^T \hat{\lambda} - \tau \sum_{i=1}^n \log(x_i) \\ &= x^T \hat{s} + b^T \hat{\lambda} - \tau \sum_{i=1}^n \log(x_i) \end{aligned}$$

Mit $B_\tau(x) \leq B_\tau(\hat{x})$ folgt $x^T \hat{s} b^T \hat{\lambda} - \tau \sum_{i=1}^n \log(x_i) \leq B_\tau(\hat{x})$, also $\sum_{i=1}^n (\hat{s}_i x_i - \tau \log(x_i)) \leq B_\tau(\hat{x}) - b^T \hat{\lambda} =: \kappa$.

Die Funktionen $x_i \mapsto (\hat{s}_i x_i - \tau \log(x_i))$ sind nach wiederum nach unten beschränkt. Daher liegen auch alle $\tilde{x} \in L_\tau$ über Null, obwohl lediglich $x \geq 0$ gefordert wurde.

L_τ ist also abgeschlossen und beschränkt \Rightarrow kompakt. B_τ ist stetig, nimmt also auf L_τ ein Minimum an. $\min B_\tau$ u.d.N. $x \in L_\tau$ besitzt also eine Lösung x_τ . Aus der Definition von L_τ folgt, dass diese Lösung schon Lösung des primalen Barriereproblems ist. \square

Satz 1.11 (Existenz des zentralen Pfades) *Ist die strikt zulässige Menge F^0 nichtleer, so besitzen die zentralen Pfad-Bedingungen für jedes $\tau > 0$ eine Lösung $(x_\tau, \lambda_\tau, s_\tau)$. Dabei sind x_τ und s_τ eindeutig bestimmt. Ist ausserdem noch A regulär, ist auch λ_τ eindeutig bestimmt.*

Beweis: Da F^0 n.V. nichtleer ist, besitzt das primale Barriere-Problem für jedes $\tau > 0$ eine Lösung x_τ . Daher besitzen auch die zentralen Pfad-Bedingungen eine Lösung $(x_\tau, \lambda_\tau, s_\tau)$, wobei die x-Komponente mit x_i zusammenfällt. Da das primale Barriere-Problem offenbar strikt konvex ist, ist x_τ eindeutig bestimmt. Mit $x_i s_i = \tau$ für $i = 1, \dots, n$ ist auch s_i eindeutig bestimmt. λ_τ ist im Allgemeinen gilt das nicht für λ , hat A aber vollen Rang, ist λ durch $A^T \lambda + s = c$ festgelegt. \square

1.4 Wiederholung: Das Newton-Verfahren

Sei $F : \mathbb{R}^p \mapsto \mathbb{R}^p$ eine stetig diffbare Funktion. Sei w^* eine Lösung des nichtlinearen Gleichungssystems $F(w) = 0$.

Bezeichnet w^k eine Näherung für w^* , so bestimmt das Newton Verfahren die nächste Iterierte w^{k+1} als Lösung des Gleichungssystems $F_k(w) = 0$, wobei $F_k : \mathbb{R}^p \mapsto \mathbb{R}^p$ die Linearisierung von F um w^k sei, d.h.

$$F_k(w) := F(w^k) + F'(w^k)(w - w^k).$$

Mit diesem F_k lautet die Berechnungsvorschrift für w^{k+1}

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

Zur praktischen Durchführung setzt man $\Delta w^k := w^{k+1} - w^k$ und bestimmt Δw^k als Lösung des LGS

$$F'(w^k)\Delta w^k = -F(w^k)$$

Dieses LGS wird auch Newton Gleichung genannt.

2 Innere-Punkte-Verfahren

Die Idee der Innere-Punkte-Verfahren ist es, das Newton Verfahren auf die zentralen Pfad-Bedingungen anzuwenden. Zu diesem Zweck setzt man

$$F_\tau(w) := F_\tau(x, \lambda, s) := \begin{pmatrix} A^T \lambda + s - c \\ Ax - b \\ XSe - \tau e \end{pmatrix},$$

mit $X := \text{diag}(x_1, \dots, x_n)$, $S := \text{diag}(s_1, \dots, s_n)$, $e = (1, \dots, 1)^T$. Mit F_τ lauten die zentralen Pfad-Bedingungen also

$$F_\tau(x, \lambda, s) = 0, \quad x > 0, \quad s > 0$$

Definition 2.1 (gewichtete Dualitätslücke) $\mu := \frac{x^T s}{n}$

Satz 2.2 Sei $w = (x, \lambda, s)$ ein gegebener Vektor mit $x > 0$ und $s > 0$. A habe vollen Rang. Dann ist $F'_\tau(x, \lambda, s)$ für jedes $\tau > 0$ regulär.

Beweis: Es gilt

$$F'_\tau(x, \lambda, s) = \begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix}$$

Sei $p = (p^1, p^2, p^3) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n$ ein beliebiger Vektor mit $F'_\tau(x, \lambda, s)p = 0$. Mit der Definition von F'_τ kann man auch schreiben:

$$A^T p^2 + p^3 = 0 \quad (I)$$

$$A p^1 = 0 \quad (II)$$

$$S p^1 + X p^3 = 0 \quad (III)$$

Multipliziert man (I) von links mit $(p^1)^T$, ergibt sich

$$0 = (p^1)^T A^T p^2 + (p^1)^T p^3 = (Ap^1)^T p^2 + (p^1)^T p^3 = (p^1)^T p^3.$$

Mit $p^3 = -X^{-1}Sp^1(III)$ folgt $(p^1)^T X^{-1}Sp^1 = 0$. Da $X^{-1}S$ positiv definit ist, folgt $p^1 = 0$, und somit auch $p^3 = 0$.

Da n.V. vollen Rang besitzt, folgt mit (I) auch $p^2 = 0$. Damit folgt die Behauptung. \square

Somit kann das Newton Verfahren auf das $F_\tau(w) = 0$ angewandt werden.

Bezeichnet $w^k = (x^k, \lambda^k, s^k)$ einen gegebenen Iterationsvektor, so setzt man $w^{k+1} = w^k + t_k \Delta w^k$ für eine gewisse Schrittweite $t_k > 0$ und einen Korrekturvektor Δw^k , der sich aus der Lösung der Newton-Gleichung

$$F'_{\tau_k}(w^k) \Delta w^k = -F_{\tau_k}(w^k)$$

ergibt. Setzt man die Definition von $F_\tau(w)$ ein, erhält man

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x^k \\ \Delta \lambda^k \\ \Delta s^k \end{pmatrix} = \begin{pmatrix} -A^T \lambda^k - s^k + c \\ -Ax^k + b \\ -X^k S^k e + \tau_k e \end{pmatrix}$$

Genügt die k -te Iterierte den beiden linearen Gleichungssystemen, die in den zentralen Pfad-Bedingungen auftreten, so sind die ersten beiden Zeilen der rechten Seite Null.

Mit $w^{k+1} = w^k + t_k \Delta w^k$ genügt dann auch die $k+1$ -te Iterierte diesen Bedingungen: Aus der ersten Blockzeile der Matrix folgt $A^T \Delta \lambda^k + \Delta s^k = 0$, aus der zweiten Zeile folgt $A \Delta x^k = 0$.

Wählt man also einen zulässigen Vektor als Startvektor, kann man die beiden ersten Zeilen der rechten Seite für die gesamte Iteration auf Null setzen.

So ergibt sich ein allgemeines Innere-Punkte-Verfahren.

(S.0) Wähle $w^0 := (x^0, \lambda^0, s^0) \in F^0$, $\epsilon \in (0, 1)$, und setze $k := 0$.

(S.1) Ist $\mu_k := (x^k)^T s^k / n \leq \epsilon$: STOP

(S.2) Wähle $\sigma_k \in [0, 1]$, und bestimme eine Lösung $\Delta w^k := (x^k, \lambda^k, s^k)$ des linearen Gleichungssystems

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -X^k S^k e + \sigma_k \mu_k e \end{pmatrix}$$

(S.3) Setze $w^{k+1} := w^k + t_k \Delta w^k$, $k \leftarrow k + 1$, und gehe zu (S.1). Dabei ist $t_k > 0$ eine Schrittweite.

2.1 Ein zulässiges Verfahren

Definiere eine Umgebung des zentralen Pfades,

$$N_{-\infty}(\gamma) := \{(x, \lambda, s) \in F^0 \mid x_i s_i \geq \gamma \mu \text{ für alle } i = 1, \dots, n\},$$

wobei $\gamma \in (0, 1)$ gegebener Parameter ist, und μ wieder die gewichtete Dualitätslücke bezeichnet.

Für $\gamma = 0$ erhalte man die Bedingung $x_i s_i \geq 0$, die ohnehin schon von F^0 gestellt wurde. Im anderen Extremfall $\gamma = 1$ hätte man eine Einschränkung auf $x_1 s_1 = \dots = x_n s_n$. Je grösser also γ ist, desto weniger Abweichungen der Komponenten voneinander ist zugelassen, und die Menge dadurch eingeschränkt. Insbesondere liegt die Lösung für jedes γ in der Menge.

(S.0) Wähle $\gamma \in (0, 1)$, $0 < \sigma_{min} < \sigma_{max} < 1$, $\epsilon \in (0, 1)$, $w^0 := (x^0, \lambda^0, s^0) \in N_{-\infty}(\gamma)$, und setze $k := 0$.

(S.1) Ist $\mu_k := (x^k)^T s^k / n \leq \epsilon$: STOP

(S.2) Wähle $\sigma_k \in [\sigma_{min}, \sigma_{max}]$, und bestimme eine Lösung $\Delta w^k := (x^k, \lambda^k, s^k)$ des linearen Gleichungssystems

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -X^k S^k e + \sigma_k \mu_k e \end{pmatrix}.$$

Sei t_k die größte Schrittweite $t \in [0, 1]$ mit

$$(x^k(t), \lambda^k(t), s^k(t)) \in N_{-\infty}(\gamma)$$

(S.3) Setze $w^{k+1} := w^k + t_k \Delta w^k$, $k \leftarrow k + 1$, und gehe zu (S.1).

3 Quellcode

Hier der Quellcode des Octave Programms, ohne Preconditions:

```
function ret = interiorpoint_feasible( A, x, lambda, s, gamma, \
                                     sigma_min, sigma_max, epsilon )

    ## define w
    w = [ x; lambda; s ];

    ## prepare newton method iteration matrix
    M = [ zeros(n,n), transpose(A), eye(n,n);
          A, zeros(m,m+n); zeros(n,n+n+m) ];

    ## iterate
    for k = 1 : 1000

        ## extract x and s
        x = w(1:n);
        s = w(n+m+1:n+m+n);

        ## compute weighed duality gap
        mu = dot(x,s) / n;
```

```

## check termination condition
if ( mu < epsilon )
    ret = x;
    return;
endif

## free choice of centering parameter (in [sigma_min,sigma_max] )
sigma = ( sigma_max - sigma_min ) * rand() + sigma_min;

S = diag(s);
X = diag(x);

## finish newton method iteration matrix
M( n+m+1:n+m+n, 1:n ) = S;
M( n+m+1:n+m+n, n+m+1:n+m+n ) = X;
b = [ zeros(n+m,1); -X * S * ones(n,1) + sigma * mu * ones(n,1) ];

## solve the above linear equation system
dw = M \ b

## use maximal stepwidth
t = 2^(3/2) * gamma * sigma / n * ( 1 - gamma ) / ( 1 + gamma )

## iteration
w = w + t * dw;

endfor

## if control flow gets here, an error occured:
error( "too many iterations" );

endfunction

```

Literatur

- [1] C. GEIGER, C. KANZOW. Theorie und Optimierung restringierter Optimierungsaufgaben. Springer, New York, Berlin, Heidelberg, 2000.