Numerikus módszerek és optimalizálás

1. előadás

Bevezetés

A direkt kereső eljárásoknál egy, az x^* minimumhelyet biztosan tartalmazó intervallumból indulunk ki. Ezt az úgynevezett befoglaló intervallumot (vagy más néven bizonytalansági intervallumot, mivel csak azt tudjuk, hogy a minimumpont itt keresendő, de a helyéről nincs információnk) szűkítjük az eljárások során lépésről lépésre.
Ehhez nem teszünk mást, mint az intervallumot részekre osztjuk (általában két belső pont felvételével) és az osztópontokban a függvényértéket kiszámítjuk, a függvényértékek nagysága alapján fogjuk az új (szűkebb) bizonytalansági intervallumot meghatározni. Ahhoz, hogy az eljárás konvergens legyen, a függvénynek bizonyos konvexitási, vagy más tulajdonsággal kell rendelkeznie.

Megjegyezzük, hogy sok esetben anélkül is alkalmazzuk az alábbi eljárásokat, hogy meggyőződnénk a konvexitásról, ekkor azonban nincs garancia az eljárások konvergenciájára.

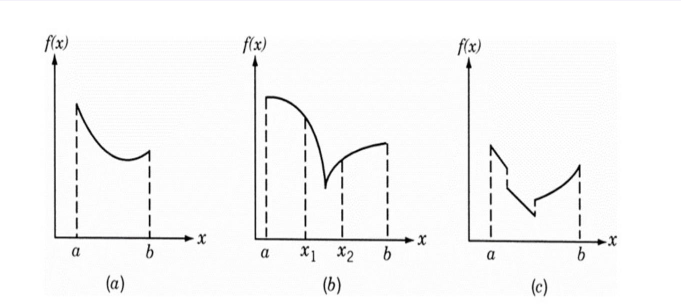
Definíció

Az $f \in C[a,b]$ függvényt unimodálisnak nevezzük, ha tetszőleges $x_1, x_2 \in [a,b]$, $x_1 < x_2$ esetén

$$\begin{aligned} x^* &< x_2 \Rightarrow f(x_2) > f(x^*) \\ x_1 &< x^* \Rightarrow f(x_1) > f(x^*) \end{aligned}$$

minden $x_1, x_2 \in [a,b]$, $x_1 < x_2$ esetén.

Feltétel nélküli optimalizálás, egyváltozós eset



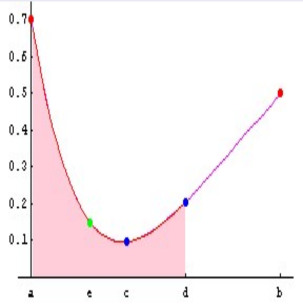
Az unimodális f függvény az x^* minimumhelytől balra szigorúan monoton csökkenő, tőle jobbra pedig szigorúan monoton növő. A szigorúan konvex és a szigorúan kvázikonvex függvények unimodálisak. Az unimodális függvények kereső eljárásai a következő észrevételen alapulnak.

Tétel

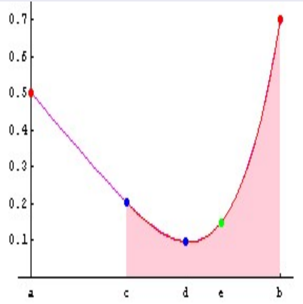
Legyen $f \in C[a, b]$ unimodális függvény és $a < c < d < b$. Ekkor

- Ha $f(c) < f(d)$, akkor $x^* \in [a, d]$.
- Ha $f(c) > f(d)$, akkor $x^* \in [c, b]$.
- Ha $f(c) = f(d)$, akkor $x^* \in [c, d]$.

Feltétel nélküli optimalizálás, egyváltozós eset



1. eset



2. eset

Megjegyezzük, hogy csak egy belső pontból származó információ alapján nem lehet a minimumhelyet tartalmazó intervallumot szűkíteni. A fenti tétel alapján a következő eljárást definiálhatjuk unimodális függvények minimumhelyének

Általános minimum kereső algoritmus

Input $[a_1, b_1]$
FOR $k = 1, 2, \dots$
    Legyen $c_k, d_k \in (a_k, b_k)$ és $c_k < d_k$.
    IF $f(c_k) < f(d_k)$
        THEN $a_{k+1} = a_k$, $b_{k+1} = d_k$
    IF $f(c_k) > f(d_k)$
        THEN $a_{k+1} = c_k$, $b_{k+1} = b_k$
    IF $f(c_k) = f(d_k)$
        THEN $a_{k+1} = c_k$, $b_{k+1} = d_k$

A kapott

$$[a_1, b_1] \supset \ldots \supset [a_k, b_k] \supset \ldots$$

intervallumsorozat közös pontként tartalmazza az x^* minimumhelyet. Ha $b_k - a_k \to 0$, akkor a minimumhely adott $\varepsilon > 0$ pontosságú közelítését véges iterációs lépésben megkaphatjuk. A c_k vagy d_k pont ismételt felhasználásával az f függvény behelyettesítéseinek számát csökkenthetjük.

A Dichotomous keresés

Legegyszerűbben úgy tudjuk biztosítani az új bizonytalansági intervallumok hosszának megegyezését, ha az intervallum középpontjától balra és jobbra azonos távolságra választjuk meg a pontokat. Jelölje a középponttól való távolságot egy alkalmasan választott $\delta > 0$ szám.
Ekkor:

$$\begin{aligned} c_k &= \frac{a_k + b_k}{2} - \delta = a_k + \frac{L_k}{2} - \delta, \\ d_k &= \frac{a_k + b_k}{2} + \delta = a_k + \frac{L_k}{2} + \delta. \qquad k = 1, 2, 3, \ldots \end{aligned}$$

Az új bizonytalansági intervallum számítása:

- Ha $f(c_k) > f(d_k)$, akkor az új bizonytalansági intervallum $[a_{k+1}, b_{k+1}] = [c_k, b_k]$.
- Ha $f(c_k) \leq f(d_k)$, akkor az új bizonytalansági intervallum $[a_{k+1}, b_{k+1}] = [a_k, d_k]$.

Nyilvánvaló, hogy mindkét esetben $L_{k+1} = \frac{L_k}{2} + \delta$.
Az eljárást addig folytatjuk, amig valamely k -ra $L_k < 2\varepsilon$, ahol $\varepsilon > 0$ a pontossági előírás. Ekkor, ha a minimumpontot a megállásnál kapott $[a_k, b_k]$ intervallum középpontjának választjuk, akkor ε -nál kisebb hibát követünk el a minimumpont közelítésében.

A dichotomous eljárás algoritmusa

Input $[a_1, b_1]$ and $\delta, \varepsilon > 0$
WHILE $b_k - a_k \geq 2\varepsilon$
  Legyen $c_k = \frac{a_k+b_k}{2} - \delta = a_k + \frac{L_k}{2} - \delta,$

  és $d_k = \frac{a_k+b_k}{2} + \delta = a_k + \frac{L_k}{2} + \delta.$

  IF $f(c_k) > f(d_k)$
    THEN $a_{k+1} = c_k, b_{k+1} = b_k$
    ELSE $a_{k+1} = a_k, b_{k+1} = c_k$
  k=k+1

Példa

Oldjuk meg az $f(x) = x^2 - 7x + 12 \to \mathtt{min!}$ egyváltozós optimalizálási feladatot Dichotomous módszerrel, ha a bizonytalansági intervallum $[a, b] = [2, 4]\,, \delta = 0.3, \varepsilon = 0.4$.

Példa

Oldjuk meg az $f(x) = x^2 - 7x + 12 \to \mathtt{min!}$ egyváltozós optimalizálási feladatot Dichotomous módszerrel, ha a bizonytalansági intervallum $[a, b] = [2, 4]$, $\delta = 0.3, \varepsilon = 0.4$.

Megoldás

1. lépés Az első közelítésbeli bizonytalansági intervallum $[a_1, b_1] = [2, 4]$, melynek hossza $L_1 = 2$. A két közbülső pont és a hozzájuk tartozó függvényértékek:

$$\begin{aligned} c_1 &= a_1 + \frac{L_1}{2} - \delta = 2.7, \qquad & f(c_1) = 0.39, \\ d_1 &= a_1 + \frac{L_1}{2} + \delta = 3.3, \qquad & f(d_1) = -0.21 \end{aligned}$$

Megoldás

2. lépés Mivel $f(c_1) > f(d_1)$, így az új bizonytalansági intervallum $[a_2, b_2] = [2.7, 4]$, melynek hossza $L_2 = 1.3$. A két közbülső pont és a hozzájuk tartozó függvényértékek:

$$\begin{aligned} c_2 &= 3.05, & f(c_2) = -0.0475, \\ d_2 &= 3.65, & f(d_2) = -0.2275 \end{aligned}$$

3. lépés Mivel $f(c_2) > f(d_2)$, így az új bizonytalansági intervallum $[a_3, b_3] = [3.05, 4]$, melynek hossza $L_3 = 0.95$. A két közbülső pont és a hozzájuk tartozó függvényértékek:

$$\begin{aligned} c_3 &= 3.225, & f(c_3) = -0.174375, \\ d_3 &= 3.825, & f(d_3) = -0.144375 \end{aligned}$$

Megoldás

4. lépés Mivel $f(c_3) < f(d_3)$, így az új bizonytalansági intervallum $[a_4, b_4] = [3.05, 3.825]$, melynek hossza $L_4 = 0.775$. Itt megállunk, mivel $L_4 < 2\varepsilon$. A minimumpont közelítése $x_{\min} \approx 3.4375$.
Általában kisebb ε értéket szokás megadni, de ezzel is elég közel kerültünk a pontos minimumhelyhez, az $\bar{x} = 3.5$ értékhez.

Az aranymetszéses keresés

Ez az eljárás egyike a leggyakrabban használtaknak. Előnye, hogy minden lépésben csak egy olyan új pont adódik, ahol ki kell kiszámolni a függvényértéket.

Legyen $\tau = \frac{\sqrt{5}-1}{2} \approx 0.618$.

Ha az aranymetszéses keresés (Golden section) eljárás befejeződik, akkor az x^* lokális minimumhely a legfeljebb 2ε hosszúságú $[a_k, b_k]$ intervallumban van. Az intervallum bármelyik pontja választható az x^* közelítéseként. Célszerű azonban az $\widetilde{x} = \frac{a_k+b_k}{2}$ közelítést alkalmazni, mert ekkor biztosan teljesül az $\left|\widetilde{x} - x^*\right| < \varepsilon$ feltétel.

Az aranymetszéses keresés eljárása

INPUT $[a_1, b_1], \varepsilon > 0, k = 1$
$c_1 = a_1 + (1-\tau)(b_1 - a_1)\,, \quad F_c = f(c_1)$
$d_1 = b_1 - (1-\tau)(b_1 - a_1)\,, \quad F_d = f(d_1)$
WHILE $b_k - a_k \geq 2\varepsilon$
    IF $F_c < F_d$
        THEN $a_{k+1} = a_k,\ b_{k+1} = d_k,\ d_{k+1} = c_k$
            $c_{k+1} = a_{k+1} + (1-\tau)(b_{k+1} - a_{k+1})$
            $F_d = F_c, \quad F_c = f(c_{k+1})$
        ELSE $a_{k+1} = c_k,\ b_{k+1} = b_k,\ c_{k+1} = d_k$
            $d_{k+1} = b_{k+1} + (1-\tau)(b_{k+1} - a_{k+1})$
            $F_c = F_d, \quad F_d = f(d_{k+1})$
    $k = k + 1$

Az aranymetsző keresés sebessége lineáris. Ugyanis az $[a_{k+1}, b_{k+1}]$ intervallum hossza pontosan τ -szorosa (0.618-szorosa) az $[a_k, b_k]$ intervallum hosszának. Jelöljük el L_k -val az $[a_k, b_k]$ hosszát, azaz $L_k = b_k - a_k$. Ekkor

$$L_{k+1} = \tau L_k = \tau^2 L_{k-1} = \cdots = \tau^k L_1.$$

Ezért adott $[a_1, b_1]$ intervallum és adott ε ismeretében előre meg lehet határozni, hogy hány lépést kell végrehajtanunk ahhoz, hogy L_{k+1} kisebb legyen ε -nál. Például, ha $L_1 = 1$ és $\varepsilon = 10^{-6}$, akkor a $\tau^k < 2 \cdot 10^{-6}$ egyenletet kell megoldanunk. Innen $k > \frac{\log 2 \cdot 10^{-6}}{\log \tau} \approx 27.2694$, azaz 28 intervallumot kell meghatározni.

Példa

Oldjuk meg az $f(x) = x^2 - 7x + 12 \to \min!$ egyváltozós optimalizálási feladatot aranymetszés módszerrel, ha a bizonytalansági intervallum $[a, b] = [2, 4]$, $\varepsilon = 0.3$.

Példa

Oldjuk meg az $f(x) = x^2 - 7x + 12 \to \min!$ egyváltozós optimalizálási feladatot aranymetszés módszerrel, ha a bizonytalansági intervallum $[a, b] = [2, 4]$, $\varepsilon = 0.3$.

Megoldás

1. lépés Az első közelítésbeli bizonytalansági intervallum $[a_1, b_1] = [2, 4]$, melynek hossza $L_1 = 2$. A két közbülső pont és a hozzájuk tartozó függvényértékek:

$$\begin{aligned} c_1 &= a_1 + 0.382 L_1 = 2.764, \qquad & f(c_1) = 0.291696, \\ d_1 &= a_1 + 0.618 L_1 = 3.236, \qquad & f(d_1) = -0.180304 \end{aligned}$$

Megoldás

2. lépés Mivel $f(c_1) > f(d_1)$, így az új bizonytalansági intervallum $[a_2, b_2] = [2.764, 4]$, melynek hossza $L_2 = 1.236$. A két közbülső pont és a hozzájuk tartozó függvényértékek az alábbiak. Ne feledjük, hogy most a c_2 pont az előző intervallumbeli d_1 értékkel azonos.

$$\begin{aligned} c_2 &= d_1 = 3.236, & f(c_2) &= -0.180304, \\ d_2 &= a_2 + 0.382L_2 = 3.527848, & f(d_2) &= -0.24922449 \end{aligned}$$

3. lépés Mivel $f(c_2) > f(d_2)$, így az új bizonytalansági intervallum $[a_3, b_3] = [3.236, 4]$, melynek hossza $L_3 = 0.764$. A két közbülső pont és a hozzájuk tartozó függvényértékek az alábbiak. Ez esetben a c_3 pont lesz az előző intervallumbeli d_2 értékkel azonos.

Megoldás

$$\begin{aligned} c_3 &= d_2 = 3.527848, & f(c_3) = -0.24922449, \\ d_3 &= a_3 + 0.618L_3 = 3.70806, & f(d_3) = -0.20671104 \end{aligned}$$

4. lépés Mivel $f(c_3) < f(d_3)$, így az új bizonytalansági intervallum $[a_4, b_4] = [3.236, 3.70806]$, melynek hossza $L_4 = 0.47206$. Itt megállunk, mivel $L_4 < 2\varepsilon$. A minimumpont közelítése $x_{\min} \approx 3.47203$.
Most még közelebb kerültünk a pontos minimumhelyhez (az $\bar{x} = 3.5$) értékhez, mint a Dichotomous keresés esetén.