Numerikus módszerek 1.

9. előadás: Gauss–Seidel iteráció, relaxációs módszer, Richardson típusú iterációk

Krebsz Anna

ELTE IK

Tartalomjegyzék

Emlékeztető: Iterációs módszerek

Az $Ax = b$ LER megoldása érdekében alakítsuk azt át $x = Bx + c$ alakúra, és valamely $x^{(0)}$ kezdőpontból végezzük az

$$x^{(k+1)} = B \cdot x^{(k)} + c \qquad (k \in \mathbb{N}_0)$$

iterációt. A fixponttétel adja meg a sorozat képletét.
A vektorsorozat bizonyos feltételek mellett konvergál a LER megoldásához (lásd fixponttétel, elégséges feltétel, szükséges és elégséges feltétel a konvergenciára a B átmenet mátrixszal).

Volt: Banach-féle fixponttétel, Jacobi-, csillapított Jacobi-iteráció.

Megjegyzés:

- 2–3 változó: felesleges $\Rightarrow$ célja a megértés
- sok változó (100, 1000): használják

Tartalomjegyzék

Átalakítás:

$$
\begin{aligned}
Ax &= b \\
(L + D + U)x &= b \\
(L + D)x &= -Ux + b \\
x &= -(L + D)^{-1}Ux + (L + D)^{-1}b
\end{aligned}
$$

Ezek alapján az iteráció a következő.

Definíció: Gauss–Seidel-iteráció

$$
x^{(k+1)} = \underbrace{-(L + D)^{-1}U}_{B_S} \cdot x^{(k)} + \underbrace{(L + D)^{-1}b}_{c_S} = B_S \cdot x^{(k)} + c_S
$$

Eml.:

$$x^{(k+1)} = -(L+D)^{-1}U \cdot x^{(k)} + (L+D)^{-1}b$$

Írjuk fel koordinátánként! (Kiderül, hogy „helyben” számolható.)

Állítás: a Gauss–Seidel-iteráció komponensenkénti alakja

$$x_i^{(k+1)} = \frac{-1}{a_{ii}} \left(\sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} + \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} - b_i \right)$$

Biz.: Alakítsunk át, majd gondoljunk bele a mátrixszorzásba.

$$\begin{aligned} (L+D)x^{(k+1)} &= -Ux^{(k)} + b \\ Dx^{(k+1)} &= -Lx^{(k+1)} - Ux^{(k)} + b \\ x^{(k+1)} &= -D^{-1}(Lx^{(k+1)} + Ux^{(k)} - b) \quad \square \end{aligned}$$

Írjuk fel az iteráció reziduum vektoros alakját!

$$(L+D)\cdot x^{(k+1)} = -U\cdot x^{(k)} + b = ((L+D)-A)\cdot x^{(k)} + b =$$
$$= (L+D)\cdot x^{(k)} + \left(-Ax^{(k)} + b\right) = (L+D)\cdot x^{(k)} + r^{(k)}$$
$$\Rightarrow \quad x^{(k+1)} = x^{(k)} + (L+D)^{-1} r^{(k)}$$

Vezessük be az $s^{(k)} := (L+D)^{-1} r^{(k)}$ segédvektort, ezzel egy lépésünk alakja:

$$x^{(k+1)} = x^{(k)} + s^{(k)}.$$

Az új reziduum vektor:

$$r^{(k+1)} = b - Ax^{(k+1)} = b - A\big(x^{(k)} + s^{(k)}\big) = r^{(k)} - As^{(k)}.$$

Algoritmus: Gauss–Seidel-iteráció

$$r^{(0)} := b - Ax^{(0)}$$

$k = 1, \ldots,$ leállásig

$$s^{(k)} := (D+L)^{-1} r^{(k)} \text{ helyett}$$

$$(D+L)\, s^{(k)} = r^{(k)} \text{ LER mo.}$$

$$x^{(k+1)} := x^{(k)} + s^{(k)}$$

$$r^{(k+1)} := r^{(k)} - As^{(k)}$$

Tétel

Ha A szig. diag. dom. a soraira, akkor az $Ax = b$ LER-re felírt Gauss–Seidel-iterációra

$$\|B_S\|_\infty \leq \|B_J\|_\infty < 1$$

teljesül, tehát az konvergens bármely $x^{(0)}$ esetén.

Biz.: Nélkül.

Tétel

Ha A szimmetrikus és pozitív definit, akkor a Gauss–Seidel-iteráció konvergens.

Biz.: Nélkül.

Induljunk a Gauss–Seidel-iteráció következő alakjából:

$$\begin{aligned}(L+D)\cdot x &= -U\cdot x+b && /\cdot\omega\\ D\cdot x &= D\cdot x && /\cdot(1-\omega)\end{aligned}$$

A kettő súlyozott összege:

$$(D+\omega L)\cdot x = [(1-\omega)D-\omega U]\cdot x+\omega b$$

$$x=(D+\omega L)^{-1}[(1-\omega)D-\omega U]\cdot x+(D+\omega L)^{-1}\omega b$$

Ezek alapján az iteráció a következő.

Definíció: relaxált Gauss–Seidel-iteráció ω paraméterrel – $S(\omega)$

$$x^{(k+1)}=\underbrace{(D+\omega L)^{-1}[(1-\omega)D-\omega U]}_{B_{S(\omega)}}\cdot x^{(k)}+\underbrace{\omega(D+\omega L)^{-1}b}_{c_{S(\omega)}}$$

Írjuk fel koordinátánként! (Kiderül, hogy „helyben" számolható.)

Állítás: $S(\omega)$ komponensenkénti alakja

$$x_i^{(k+1)} = (1-\omega)\cdot x_i^{(k)} + \omega \cdot x_{i,S}^{(k+1)},$$

ahol $x_{i,S}^{(k+1)}$ a hagyományos Seidel-módszer $(S = S(1))$ által adott, azaz

$$x_{i,S}^{(k+1)} = \frac{-1}{a_{ii}} \left(\sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} + \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} - b_i \right).$$

Minden k. lépés az $i = 1, 2, \ldots, n$ sorrendben számolandó.

Biz.: Alakítsunk át, majd gondoljunk bele a mátrixszorzásba.

$$\begin{aligned}(D+\omega L)x^{(k+1)} &= (1-\omega)Dx^{(k)} - \omega Ux^{(k)} + \omega b\\ Dx^{(k+1)} &= (1-\omega)Dx^{(k)} - \omega Lx^{(k+1)} - \omega Ux^{(k)} + \omega b\\ x^{(k+1)} &= (1-\omega)x^{(k)} - \omega \underbrace{D^{-1}\left(Lx^{(k+1)} + Ux^{(k)} - b\right)}_{\text{Lásd } S(1)\text{-nél.}}\end{aligned}$$

A koordinátánkénti alakja:

$$x_i^{(k+1)} = (1-\omega)\cdot x_i^{(k)} - \frac{\omega}{a_{ii}}\left(\sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} + \sum_{j=i+1}^{n} a_{ij}x_j^{(k)} - b_i\right).$$

Megj.: Vigyázat! $x^{(k+1)} = (1-\omega)\cdot x^{(k)} + \omega\cdot x_S^{(k+1)}$ nem igaz (tehát az egész vektorra); csak komponensenként.

Relaxált Gauss–Seidel-iteráció

Írjuk fel az iteráció reziduum vektoros alakját!

$$(D+\omega L)\,x^{(k+1)} = ((1-\omega)D - \omega U)\cdot x^{(k)} + \omega b =$$

$$= D\,x^{(k)} + \omega(\underbrace{(-D-U)}_{L-A})\cdot x^{(k)} + \omega b =$$

$$= D\,x^{(k)} + \omega L x^{(k)} + \omega\underbrace{\left(-Ax^{(k)} + b\right)}_{r^{(k)}} = (D+\omega L)\,x^{(k)} + \omega r^{(k)}$$

$$\Rightarrow \quad x^{(k+1)} = x^{(k)} + \omega(D+\omega L)^{-1} r^{(k)}$$

Vezessük be az $s^{(k)} := \omega(D+\omega L)^{-1} r^{(k)}$ segédvektort, ezzel egy lépésünk alakja: $x^{(k+1)} = x^{(k)} + s^{(k)}$.
Az új reziduum vektor:

$$r^{(k+1)} = b - Ax^{(k+1)} = b - A(x^{(k)} + s^{(k)}) = r^{(k)} - As^{(k)}.$$

Algoritmus: relaxált Gauss–Seidel-iteráció $S(\omega)$

$$r^{(0)} := b - Ax^{(0)}$$

$k = 1, \dots,$ leállásig

$$s^{(k)} := \omega(D + \omega L)^{-1} r^{(k)} \text{ helyett}$$

$$(D + \omega L)\, s^{(k)} = \omega r^{(k)} \text{ LER mo.}$$

$$x^{(k+1)} := x^{(k)} + s^{(k)}$$

$$r^{(k+1)} := r^{(k)} - As^{(k)}$$

Relaxált Gauss–Seidel-iteráció

Tétel: a relaxált Gauss–Seidel-módszer $S(\omega)$ konvergenciájáról

Ha egy mátrixra az $S(\omega)$ módszer konvergens, akkor $0 < \omega < 2$.

Megjegyzés:

- Ha $\omega \notin (0, 2)$, akkor általában nem konvergál.
- A relaxált Seidel-módszert gyakran alkalmazzák...

Lemma

$$\det B = \prod_{i=1}^{n} \lambda_i(B)$$

Biz. lemma: Írjuk fel a B mátrix karakterisztikus polinomját, amelyről tudjuk, hogy gyökei a mátrix sajátértékei; majd rendezzük λ hatványai szerint:

$$p(\lambda) = \det(B - \lambda I) = \prod_{i=1}^{n} (\lambda_i - \lambda) = (-1)^n \cdot \lambda^n + \ldots + \prod_{i=1}^{n} \lambda_i.$$

A $\lambda = 0$ értéket behelyettesítve a konstans tagot kapjuk, amire:

$$p(0) = \det(B) = \prod_{i=1}^{n} \lambda_i.$$

$\square$

Biz. tétel: A konvergencia ekvivalens feltételéből, azaz a

$$\varrho(B_{S(\omega)}) < 1$$

állításból kell ω kívánt becslését előállítanunk. Egyrészt

$$\varrho(B_{S(\omega)}) < 1 \;\Rightarrow\; \left|\lambda_i(B_{S(\omega)})\right| < 1 \;\Rightarrow\; \Rightarrow \left|\prod_{i=1}^{n} \lambda_i(B_{S(\omega)})\right| < 1 \;\Rightarrow\; \left|\det(B_{S(\omega)})\right| < 1.$$

Az iteráció mátrixa

$$B_{S(\omega)} = (D + \omega L)^{-1}[(1-\omega)D - \omega U].$$

Kihasználjuk, hogy háromszögmátrixok determinánsa a főátlóbeli elemek szorzata (tehát nem függ a diagonálison kívüli elemektől).

Biz. tétel folyt.:

$$\left|\det(B_{S(\omega)})\right| = \underbrace{\left|\det\left((D+\omega L)^{-1}\right)\right|}_{1/|\det(D)|} \cdot \underbrace{\left|\det\left((1-\omega)D-\omega U\right)\right|}_{|1-\omega|^n \cdot |\det(D)|} =$$

$$= \frac{1}{\det(D)} \cdot |1-\omega|^n \cdot |\det(D)| = |1-\omega|^n < 1$$

Ebből pedig $|1-\omega| < 1$ következik, ami ekvivalens a $0 < \omega < 2$ becsléssel. $\square$

Tétel: a relaxált Gauss–Seidel-módszer $S(\omega)$ konvergenciájáról

Ha az egyenletrendszer mátrixa szimmetrikus, pozitív definit és $\omega \in (0,2)$, akkor az $S(\omega)$ módszer konvergens.

Biz.: nélkül.

Tétel: $S(\omega)$ tridiagonális mátrixokra

Ha a LER mátrixa tridiagonális, akkor a Jacobi- és Gauss–Seidel-iteráció egyszerre konvergens vagy divergens

$$\text{azaz } \varrho(B_S) = \varrho(B_J)^2.$$

Ez azt jelenti, hogy konvergencia esetén a Gauss–Seidel-iteráció kétszer gyorsabb,

Biz.: nélkül.

Tétel: $S(\omega)$ szimmetrikus, pozitív definit és tridiagonális mátrixokra

Ha a LER mátrixa tridiagonális, szimmetrikus és pozitív definit, akkor a Jacobi-, Gauss–Seidel- és relaxált Gauss–Seidel-iteráció is konvergens. Megadható $S(\omega)$-ra optimális paraméter

$$\omega_0 = \frac{2}{1 + \sqrt{1 - \varrho(B_J)^2}}.$$

Továbbá,

- ha $\varrho(B_J) = 0$, akkor $\omega_0 = 1$ és $\varrho(B_S) = \varrho(B_{S(\omega_0)}) = 0$,
- $\varrho(B_J) \neq 0$, akkor $\varrho(B_{S(\omega_0)}) = \omega_0 - 1 < \varrho(B_S) = \varrho(B_J)^2$.

Biz.: nélkül.

Megj.:

- Az utóbbi két tétel blokktridiagonális mátrixokra is igaz, a megfelelő blokkiterációkra.
- Az iterációs módszer konvergencia sebessége a q kontrakciós együtthatótól függ. Minél közelebb van 0-hoz, annál gyorsabb a módszer, míg, ha 1-hez van közel, akkor nagyon lassú. A kontrakciós együtthatót $q = \|B\|$-ként kapjuk.
- Mivel bármely normára $\inf\{\|B\| : B \text{ indukált norma}\} = \varrho(B)$, ezért a spektrálsugár határozza meg a konvergencia sebességét.

Példa

Mit állíthatunk a következő mátrixra felírt Jacobi- és Gauss–Seidel-iterációk konvergenciájáról?

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

A szimmetrikus, pozitív definit és tridiagonális, alkalmazhatók rá a tanult tételek:

- A $J(1)$ iteráció konvergens minden kezdővektorra.
- Ha $\omega \in (0; 1)$-re, akkor $J(\omega)$ iteráció konvergens minden kezdővektorra.
- Az $S(1)$ iteráció konvergens minden kezdővektorra.
- Az $S(\omega)$ iteráció konvergens minden kezdővektorra pontosan az $\omega \in (0; 2)$ értékekre.

Példa

$$B_J = \frac{1}{2}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad B_S = \frac{1}{8}\begin{bmatrix} 0 & 4 & 0 \\ 0 & 2 & 4 \\ 0 & 1 & 2 \end{bmatrix}$$

B_J sajátértékei: $0, \pm\frac{1}{\sqrt{2}}$, így $\varrho(B_J) = \frac{1}{\sqrt{2}}$.

B_S sajátértékei: $0, 0, \frac{1}{2}$, így $\varrho(B_S) = \frac{1}{2}$.

$S(\omega)$-ra az optimális paraméter:

$$\omega_0 = \frac{2}{1 + \sqrt{1 - \left(\frac{1}{\sqrt{2}}\right)^2}} = \frac{2}{1 + \frac{1}{\sqrt{2}}} = \frac{2\sqrt{2}}{1 + \sqrt{2}} = 4 - 2\sqrt{2} \approx 1,1716\ldots$$

$\varrho(B_{S(\omega_0)}) = \omega_0 - 1 = 3 - 2\sqrt{2} \approx 0,1716\ldots$
Nézzük meg a LER-re a csillapított Jacobi- és a relaxált Gauss–Seidel-iteráció vizsgálatát Maple-ben.

Richardson-iteráció

Tekintsük az $Ax = b$ LER-t, ahol A szimmetrikus, pozitív definit mátrix és $p \in \mathbb{R}$.

$$\begin{aligned}
Ax &= b \\
p \cdot Ax &= p \cdot b \\
0 &= -pAx + pb \\
x &= x - pAx + pb = (I - pA)x + pb
\end{aligned}$$

Ezek alapján az iteráció a következő.

Definíció: Richardson-iteráció p paraméterrel – $R(p)$

$$x^{(k+1)} = \underbrace{(I - pA)}_{B_{R(p)}} \cdot x^{(k)} + \underbrace{pb}_{c_{R(p)}} = B_{R(p)} \cdot x^{(k)} + c_{R(p)}$$

Írjuk fel az iteráció reziduum vektoros alakját!

$$x^{(k+1)} = x^{(k)} - pAx^{(k)} + pb = x^{(k)} + p \cdot \left(-Ax^{(k)} + b\right) =$$
$$= x^{(k)} + pr^{(k)}$$

Vezessük be az $s^{(k)} := pr^{(k)}$ segédvektort, ezzel egy lépésünk alakja:

$$x^{(k+1)} = x^{(k)} + s^{(k)}.$$

Az új reziduum vektor:

$$r^{(k+1)} = b - Ax^{(k+1)} = b - A(x^{(k)} + s^{(k)}) = r^{(k)} - As^{(k)}.$$

Algoritmus: Richardson-iteráció

$$r^{(0)} := b - Ax^{(0)}$$

$k = 1, \ldots,$ leállásig

$$s^{(k)} := pr^{(k)}$$

$$x^{(k+1)} := x^{(k)} + s^{(k)}$$

$$r^{(k+1)} := r^{(k)} - As^{(k)}$$

Megjegyzés: Érdemes meggondolni, hogy ha az $Ax = b$ helyett a $D = \operatorname{diag}(a_{11}, \ldots, a_{nn})$ diagonális mátrix-szal a $D^{-1}Ax = D^{-1}b$ LER-re alkalmazzuk az $R(p)$ iterációt, akkor az eredeti LER-re felírt $J(p)$ csillapított Jacobi-iterációt kapjuk.

Tétel: A Richardson-iteráció konvergenciája

Ha az $A \in \mathbb{R}^{n\times n}$ mátrix szimmetrikus, pozitív definit és sajátértékeire $m = \lambda_1 \leq \cdots \leq \lambda_n = M$ teljesül, akkor $R(p)$ (azaz az $Ax = b$ LER-re felírt $p \in \mathbb{R}$ paraméterű Richardson-iteráció) pontosan a

$$p \in \left(0, \frac{2}{M}\right),$$

paraméter értékekre konvergens minden kezdővektor esetén. Az optimális paraméter $p_0 = \frac{2}{M+m}$, a hozzá kapcsolódó kontrakciós együttható pedig:

$$\varrho(B_{R(p_0)}) := \frac{M-m}{M+m} = \left\|B_{R(p_0)}\right\|_2 = q.$$

Bizonyítás:

1. $B_{R(p)}$ sajátértékei: $\lambda_i(p) = 1 - p \cdot \lambda_i$, hiszen

$$Av = \lambda_i v \quad \Rightarrow \quad (I - pA)v = v - pAv = v - p\lambda_i v = (1 - p\lambda_i)v.$$

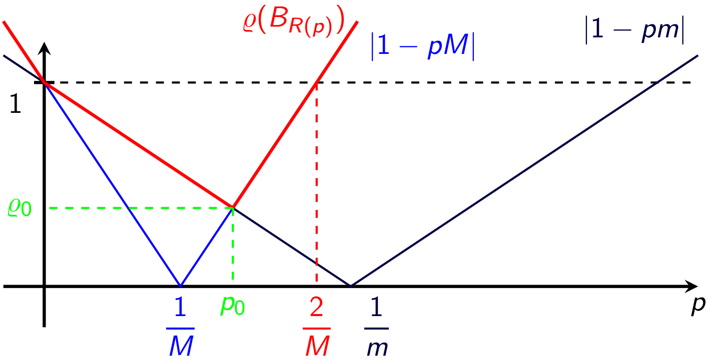
Vagyis:

$$\begin{aligned} \lambda_1(p) &= 1 - p \cdot \lambda_1 = 1 - pm, \\ \lambda_2(p) &= 1 - p \cdot \lambda_2, \\ &\vdots \\ \lambda_n(p) &= 1 - p \cdot \lambda_n = 1 - pM. \end{aligned}$$

2. $B_{R(p)}$ spektrálsugara rögzített p-re
$\varrho(B_{R(p)}) = \max_{i=1}^{n} |1 - p \cdot \lambda_i|.$

❸ Ábrázoljuk az $|1 - p \cdot \lambda_i|$ függvényeket $(i = 1, 2, \ldots, n)$! (Ezek p-től függenek.)

$$1 - p \cdot \lambda_i = 0 \quad \Longleftrightarrow \quad p = \frac{1}{\lambda_i}$$



4. $R(p)$ konvergens, ha $\varrho(B_{R(p)}) < 1$, azaz ha $p \in \left(0, \frac{2}{M}\right)$.
   Ezek az $|1 - pM| = 1$ egyenlet megoldásai.

5. Továbbá az optimális paramétert az

$$|1 - pM| = |1 - pm|$$

egyenlet megoldása adja. (Nem a 0, hanem a másik.)

$$\begin{aligned} -1 + pM &= 1 - pm \\ pM + pm &= 2 \\ p(M + m) &= 2 \quad \Longrightarrow \quad p_0 = \frac{2}{M + m} \end{aligned}$$

6

$$\varrho(B_{R(p_0)}) = 1 - p_0 \cdot m = \frac{M+m}{M+m} - \frac{2m}{M+m} = \frac{M-m}{M+m}.$$

7 Mivel A szimmetrikus, így $B_{R(p)}$ is, ezért a spektrálsugara és kettes normája megegyezik. Az eredményül kapott spektrálsugár egyben kettes normabeli kontrakciós együttható:

$$q = \frac{M-m}{M+m}.$$

□

Példa

Adjuk meg, hogy a Richardson-iteráció mely $p \in \mathbb{R}$ paraméterek mellett konvergens a következő egyenletrendszer esetén – mely ugyanaz, mint az imént. Mi az optimális paraméter és a hozzá tartozó „átmenetmátrix” spektrálsugara?

$$Ax = b, \qquad \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \cdot x = \begin{bmatrix} 1 \\ 5 \end{bmatrix}.$$

A mátrix sajátértékei 2 és 4.

$M = 4$, $m = 2$, így a $p \in (0, \frac{2}{M}) = (0; \frac{1}{2})$ értékekre a Richardson-iteráció konvergens bármely kezdővektor esetén. Az optimális paraméter

$$p_0 = \frac{2}{M+m} = \frac{2}{4+2} = \frac{1}{3}$$

és a hozzá tartozó átmenetmátrix spektrálsugara

$$\varrho(B_{R(p_0)}) = \frac{4-2}{4+2} = \frac{1}{3}.$$

Mivel A szimmetriája öröklődik $B_{R(p)}$-re, így az átmenetmátrix is szimmetrikus, így

$$\left\| B_{R(p_0)} \right\|_2 = \varrho(B_{R(p_0)}) = \frac{1}{3} = q$$

a kontrakciós együttható a kettes normában.

Példák Matlab-ban

1. A tapasztalati kontrakciós együtthatók szemléltetése a relaxációs módszer esetén.
2. A Richardson-iteráció viselkedésének vizsgálata különböző paraméterek mellett.

1. Példa:

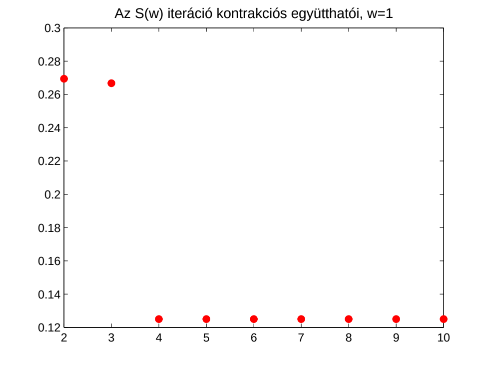
A LER alakja $Ax = b$, ahol

$$A = \begin{bmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{bmatrix}, \quad b = \begin{bmatrix} 3 \\ 2 \\ 3 \end{bmatrix}, \quad x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Vizsgáljuk a relaxációs módszer tapasztalati kontrakciós együtthatóit $\omega = 1, 0.8, 0.6, 1.033, -0.1, 2, 2.5$ esetén!

Tapasztalati kontrakciós együttható vizsgálata

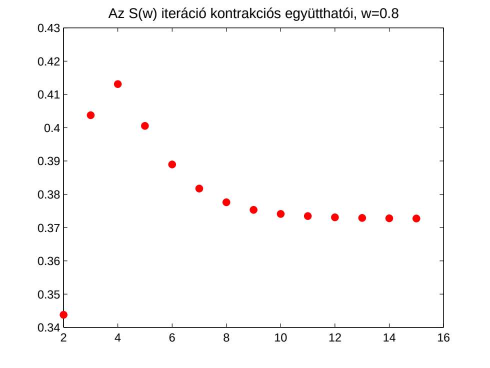
1. Példa:



$q \approx 0.1250$

Tapasztalati kontrakciós együttható vizsgálata

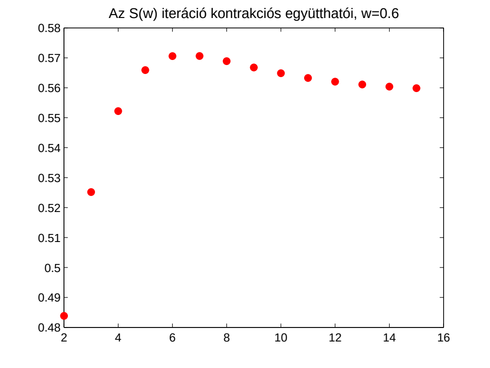
1. Példa:



$q \approx 0.3750$

Tapasztalati kontrakciós együttható vizsgálata

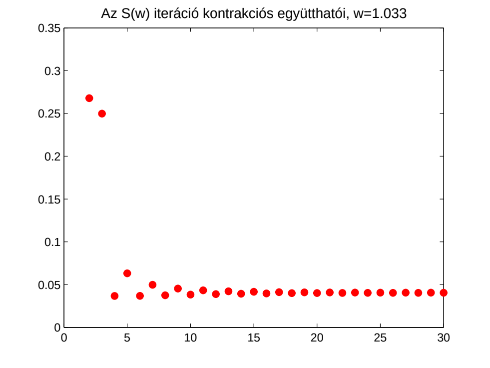
1. Példa:



$q \approx 0.5650$

Tapasztalati kontrakciós együttható vizsgálata

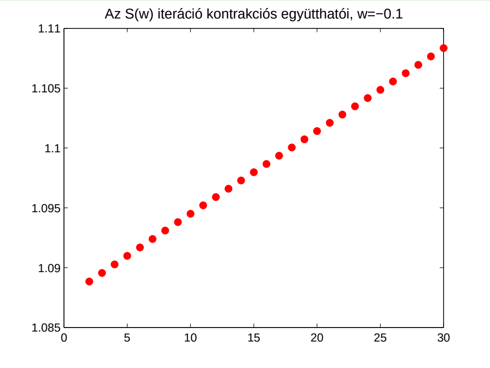
1. Példa:



$q \approx 0.0404$

Tapasztalati kontrakciós együttható vizsgálata

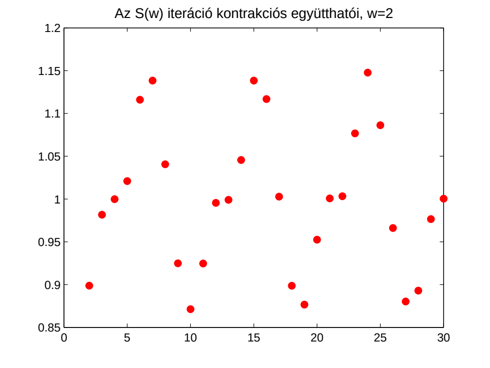
1. Példa:



$q > 1$, divergens iteráció

Tapasztalati kontrakciós együttható vizsgálata

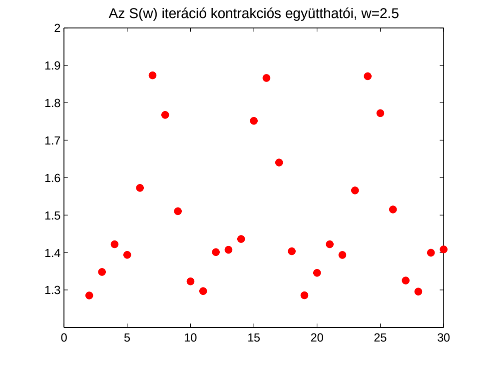
1. Példa:



$q > 1$, divergens iteráció

Tapasztalati kontrakciós együttható vizsgálata

1. Példa:



$q > 1$, divergens iteráció