

Matematika 3.

Építészeknek

Dr. Barabás Béla
egyetemi docens
BME Természettudományi Kar
Matematika Intézet
Sztochasztika Tanszék

2012. április 26.

Tartalomjegyzék

1.	Görbék illesztése adott pontokra	6
1.1.	Lagrange interpoláció	6
1.2.	Newton interpoláció	11
1.3.	Hermite interpoláció	12
1.4.	Spline	15
1.4.1.	Lineáris spline "sátorfüggvényekből"	16
1.4.2.	Lineáris spline "csonkolt" függvényekből	17
1.4.3.	Köbös spline "csonkolt" függvényekből	18
1.5.	Felületek illesztése	20
2.	Végtelen sorok	27
2.1.	Numerikus sorok	27
2.1.1.	Pozitív tagú sorok konvergencia-kritériumai	30
2.1.2.	Alternáló sorok	35
2.1.3.	Végtelen összegek tulajdonságai	35
2.2.	Függvénysorok	36
2.2.1.	Hatványsorok	40
2.2.2.	Taylor sor	41
2.2.3.	Ortogonalis sorok	43
2.2.4.	Fourier sor	45
3.	Egyenletek, egyenletrendszerek numerikus megoldása	46
3.1.	Bevezetés	46
3.2.	Intervallum felezés	47
3.3.	Iteráció	48
3.4.	Húrmódszer	48
3.5.	Érintő módszer	49
3.6.	Newton módszer	49

3.7.	Lineáris egyenletrendszerek megoldása	49
3.7.1.	Lineáris egyenletrendszerek direkt megoldása	50
3.7.2.	Lineáris egyenletrendszerek iterációs megoldása	53
3.7.3.	Relaxációs módszerek lineáris egyenletrendszerek iterációs megoldására	56
4.	Közönséges elsőrendű differenciálegyenletek numerikus megoldása	59
4.1.	Szukcesszív approximáció	61
4.2.	Euler módszer	62
4.3.	Runge - Kutta módszer	63
5.	Másod (és magasabb) rendű lineáris differenciálegyenletek	64
5.1.	Állandó együtthatós eset	68
5.2.	Magasabb rendű állandó együtthatós lineáris differenciálegyenletek . .	69
5.3.	Másodrendű lineáris differenciálegyenletek integrálása sorok segítségével	69
5.4.	Véges-elem módszer	70
5.5.	Másodrendű differenciálegyenletek mechanikai alkalmazásai	72
6.	Valószínűségszámítás alapjai	75
6.1.	A valószínűség fogalma	76
6.2.	Klasszikus valószínűségi mező	77
6.3.	Feltételes valószínűség	77
7.	Diszkrét valószínűségi változó	79
7.1.	Várható érték, szórás	79
7.2.	Nevezetes diszkrét eloszlások	80
8.	Folytonos valószínűségi változó	81
8.1.	Várható érték, szórás	82
8.2.	Nevezetes folytonos eloszlások	82
9.	Statisztika	85
9.1.	Statisztikai minta	85
9.2.	Mintavételi módszerek	85
9.3.	Empirikus eloszlásfüggvény	86
10.	Paraméterbecslés	88
10.1.	Pontbecslés	88
10.2.	Konfidencia intervallum	89
11.	Hipotézisvizsgálat	95
11.1.	u-próba, t-próba	98

11.1.1.	Egymintás próba várható értékre	98
11.1.2.	Kétmintás próba várható értékre	99
11.1.3.	Párosított t-próba	99
11.1.4.	Arány tesztelése	100
11.1.5.	Arányok egyezése	100
12.	Illeszkedésvizsgálat	100
12.1.	χ^2 próba	101
12.2.	Kolmogorov próba	101
13.	Függőség jellemzése	101
13.1.	Korreláció	101
13.2.	Regresszió	101
	Irodalomjegyzék	101

1. Görbék illesztése adott pontokra

A probléma: Olyan polinom előállítás, amelynek grafikonja adott pontokon megy át. Legyenek az adott pontok $(x_i; y_i)$ $i = 1, 2, \dots, n$, azaz a polinom értéke az x_1, x_2, \dots, x_n helyen legyen y_1, y_2, \dots, y_n .

1.1. Tétel.

Ha x_1, x_2, \dots, x_n különböző valós számok, akkor van olyan legfeljebb $n - 1$ -ed fokú

$$P_{n-1}(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \dots + a_1x + a_0$$

polinom, amelynek értéke az x_i helyen éppen y_i , vagyis $P_{n-1}(x_i) = y_i$ minden $i = 1, 2, \dots, n$, esetén.

1.2. Tétel.

A fenti tételnek eleget tevő polinom egyértelmű, azaz ha a $P_{n-1}(x)$ és $Q_{n-1}(x)$ polinomok olyanok, hogy $P_{n-1}(x_i) = Q_{n-1}(x_i)$, akkor $P_{n-1}(x) = Q_{n-1}(x)$.

1.1. Lagrange interpoláció

1.1. Definíció.

Az $l_i(x)$ Lagrange-féle i -ik interpolációs alappolinom olyan tulajdonságú, hogy az x_1, x_2, \dots, x_n helyeken 0 értéket vesz fel, kivéve az x_i helyet, ahol is a polinom értéke 1. Tehát:

$$l_i(x) = \begin{cases} 0 & \text{ha } x = x_j \quad j = 1, 2, \dots, i-1, i+1, \dots, n \\ 1 & \text{ha } x = x_i \end{cases} \quad (1)$$

Előállítás:

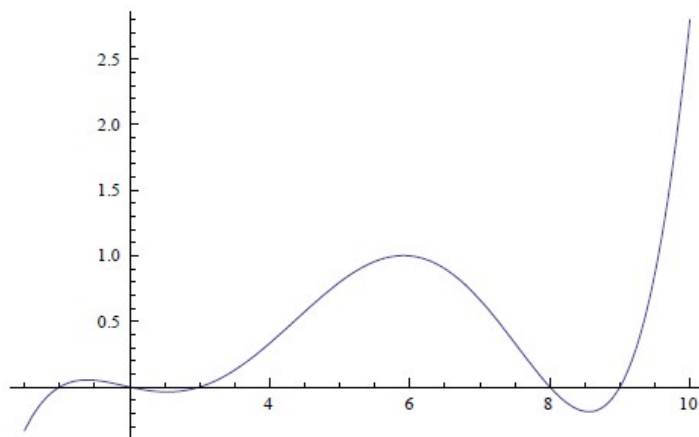
$$l_i(x) = \frac{(x - x_1)(x - x_2) \cdots (x - x_{i-1})(x - x_{i+1}) \cdots (x - x_n)}{(x_i - x_1)(x_i - x_2) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - x_n)}.$$

Ezek segítségével a kívánt polinom előállítás:

1.3. Tétel.

Lagrange-féle interpolációs polinom.

$$P(x) = \sum_{i=1}^n y_i \cdot l_i(x).$$



1. ábra. Az l_4 Lagrange alappolinom az $x = 1, 2, 3, 6, 8, 9$ alappontokra

1.1. Megjegyzés.

Szokás bevezetni az

$$\omega(x) = (x - x_1)(x - x_2) \cdots (x - x_n)$$

jelölést. Ennek segítségével

$$l_i(x) = \frac{\omega(x)}{(x - x_i) \cdot \omega'(x_i)}$$

1.2. Megjegyzés. A Lagrange interpoláció előnye.

Természetes gondolat lenne az interpolációs feladat megoldásához úgy hozzá állni, hogy a keresett $n - 1$ -ed fokú polinomnak n darab együtthatója van és van hozzá n darab feltétel, amelynek felírásakor az alábbi lineáris egyenletrendszer adódik:

$$\begin{aligned} a_{n-1}x_1^{n-1} + a_{n-2}x_1^{n-2} + \cdots + a_1x_1 + a_0 &= y_1 \\ a_{n-1}x_2^{n-1} + a_{n-2}x_2^{n-2} + \cdots + a_1x_2 + a_0 &= y_2 \\ &\vdots \\ a_{n-1}x_i^{n-1} + a_{n-2}x_i^{n-2} + \cdots + a_1x_i + a_0 &= y_i \end{aligned} \tag{2}$$

$$\begin{aligned} &\vdots \\ a_{n-1}x_n^{n-1} + a_{n-2}x_n^{n-2} + \cdots + a_1x_n + a_0 &= y_n \end{aligned} \tag{3}$$

Itt tehát az ismeretlenek az $a_{n-1}, a_{n-2}, \dots, a_0$ polinomegyütthatók. Az egyenletrendszer megoldása problémás, mert az együttható mátrix, az úgynevezett Vandermond-mátrix

$$V_n = \begin{bmatrix} x_1^{n-1} & x_1^{n-2} & \cdots & x_1 & 1 \\ x_2^{n-1} & x_2^{n-2} & \cdots & x_2 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_n^{n-1} & x_n^{n-2} & \cdots & x_n & 1 \end{bmatrix}$$

gyengén kondicionált, ezért a numerikus megoldás sok esetben rendkívül bizonytalan.

1.1. Példa.

Illesszünk interpolációs polinomot az $A(1; 3)$, $B(2; 8)$ és $C(4; 12)$ pontokra!

Megoldás: Három pontra egy legfeljebb másodfokú polinomot illesztünk. A Lagrange-féle alappolinomok:

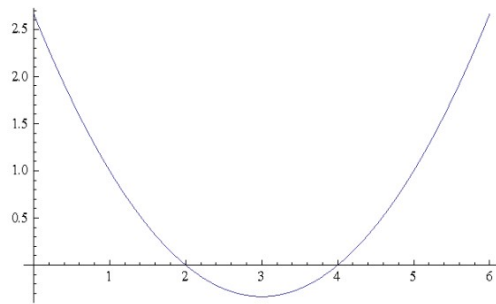
$$l_1(x) = \frac{(x-2)(x-4)}{(1-2)(1-4)} = \frac{(x-2)(x-4)}{3} = \frac{x^2 - 6x + 8}{3}$$

$$l_2(x) = \frac{(x-1)(x-4)}{(2-1)(2-4)} = \frac{(x-1)(x-4)}{-2} = \frac{x^2 - 5x + 4}{-2}$$

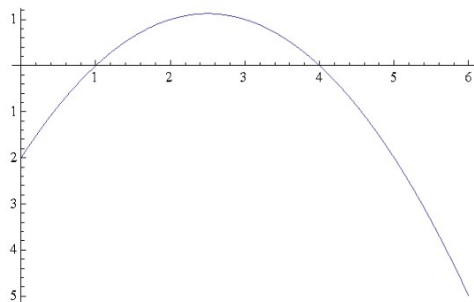
$$l_3(x) = \frac{(x-1)(x-2)}{(4-1)(4-2)} = \frac{(x-1)(x-2)}{6} = \frac{x^2 - 3x + 2}{6}$$

Ezért az interpolációs polinom:

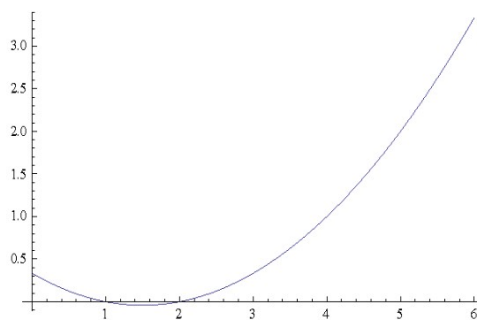
$$\begin{aligned} P(x) &= 3 \cdot l_1(x) + 8 \cdot l_2(x) + 12 \cdot l_3(x) = \\ &= x^2 - 6x + 8 - 4x^2 + 20x - 16 + 2x^2 - 6x + 4 = -x^2 + 8x - 4 \end{aligned}$$



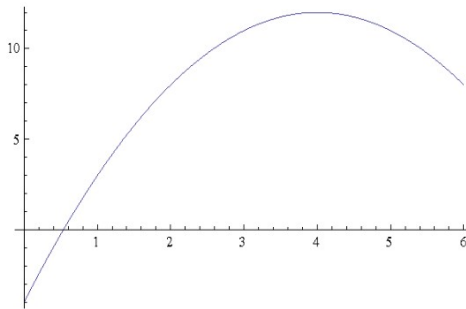
2. ábra. Az $l_1(x)$ Lagrange alappolinom az $x = 1, 2, 4$ alappontokra



3. ábra. Az $l_2(x)$ Lagrange alappolinom az $x = 1, 2, 4$ alappontokra



4. ábra. Az $l_3(x)$ Lagrange alappolinom az $x = 1, 2, 4$ alappontokra



5. ábra. A Lagrange-féle interpolációs-polinom az $x = 1, 2, 4$ alappontokra

1.2. Példa.

Illesszünk interpolációs polinomot az $A(1; 6)$, $B(2; 11)$ és $C(4; 27)$ pontokra!

Megoldás: Három pontra egy legfeljebb másodfokú polinomot illesztünk. A Lagrange-féle alappolinomok most ugyan azok, mint az előző példában, ezért a keresett interpolációs polinom:

$$\begin{aligned} P(x) &= 6 \cdot l_1(x) + 11 \cdot l_2(x) + 27 \cdot l_3(x) = \\ &= 2(x^2 - 6x + 8) - \frac{11}{2}(x^2 - 5x + 4) + \frac{9}{2}(x^2 - 3x + 2) = x^2 + 2x + 3 \end{aligned}$$

Előnye a Lagrange-féle interpolációnak, hogy ha ugyanazon beállított alappontoknál többször mérjük a függvényértékeket és mindig illeszteni kell polinomot, akkor a számításokat nem kell elölről kezdeni, mert ugyanazok az $l_i(x)$ alappolinomok használhatók.

1.3. Példa.

Illesszünk interpolációs polinomot az $A(1; 3)$, $B(2; 8)$, $C(4; 12)$ és $D(3; 7)$ pontokra!

Megoldás: Négy pontra egy legfeljebb harmadfokú polinomot illesztünk. A Lagrange-féle alappolinomok:

$$\begin{aligned} l_1(x) &= \frac{(x-2)(x-3)(x-4)}{(1-2)(1-3)(1-4)} = \frac{x^3 - 9x^2 + 26x - 24}{-6} \\ l_2(x) &= \frac{(x-1)(x-3)(x-4)}{(2-1)(2-3)(2-4)} = \frac{x^3 - 8x^2 + 19x - 12}{2} \\ l_3(x) &= \frac{(x-1)(x-2)(x-4)}{(3-1)(3-2)(3-4)} = \frac{x^3 - 7x^2 + 14x - 8}{-2} \\ l_4(x) &= \frac{(x-1)(x-2)(x-3)}{(4-1)(4-2)(4-3)} = \frac{x^3 - 6x^2 + 11x - 6}{6} \end{aligned}$$

Ezért az interpolációs polinom:

$$P(x) = 3 \cdot l_1(x) + 8 \cdot l_2(x) + 7 \cdot l_3(x) + 12 \cdot l_4(x) = 2x^3 - 15x^2 + 36x - 20$$

1.3. Megjegyzés. A Lagrange interpoláció hátránya.

A Lagrange interpoláció hátránya, hogy ha már elkészítettük az x_1, x_2, \dots, x_n helyekhez tartozó interpolációt és felmerül az igénye annak, hogy egy újabb, mondjuk x_{n+1} helyen is végezzünk mérést, akkor az egész számítást előlről kell kezdeni. Ugyanis az $x_1, x_2, \dots, x_n, x_{n+1}$ pontokhoz tartozó alappolinomok lényegesen különböznek az előzőektől. Ezt a hátrányt küszöböli ki a Newton-féle interpoláció.

1.2. Newton interpoláció

Tegyük fel, hogy az x_1, x_2, \dots, x_n helyekhez tartozó és az

$$L(x_1) = y_1, L(x_2) = y_2, \dots, L(x_n) = y_n$$

feltételt kielégítő $L(x)$ interpolációs polinomot már ismerjük. Keressük azt a $P(x)$ interpolációs polinomot, amelyik az x_1, x_2, \dots, x_n helyeken ugyanazokat az y_1, y_2, \dots, y_n értékeket veszi fel, mint a már ismert $L(x)$ polinom, és az x_{n+1} helyen y_{n+1} az értéke. Azaz $L(x)$ ismeretében keressük a

$$P(x_1) = y_1, P(x_2) = y_2, \dots, P(x_{n+1}) = y_{n+1}$$

feltételt kielégítő interpolációs polinomot.

Megoldás: Keressük a $P(x)$ polinomot

$$P(x) = L(x) + \lambda \cdot (x - x_1)(x - x_2) \cdots (x - x_n)$$

alakban. Ezzel nem rontjuk el az x_1, x_2, \dots, x_n helyeken felvett értékeket. A λ ismeretlen meghatározása egyszerű:

$$\lambda = \frac{P(x_{n+1}) - L(x_{n+1})}{(x_{n+1} - x_1)(x_{n+1} - x_2) \cdots (x_{n+1} - x_n)}.$$

1.4. Példa.

Illesszünk interpolációs polinomot az $A(1; 3)$ $B(2; 8)$ $C(4; 12)$ és $D(3; 7)$ pontokra annak ismeretében, hogy a (1.1) példában már meghatároztuk az $A(1; 3)$ $B(2; 8)$ és $C(4; 12)$ pontokra illeszkedő $L(x) = -x^2 + 8x - 4$ polinomot!

Megoldás: Felírhatjuk a

$$P(x) = L(x) + \lambda \cdot (x - 1)(x - 2)(x - 4)$$

egyenletet, amiből

$$\lambda = \frac{P(3) - L(3)}{(3 - 1)(3 - 2)(3 - 4)} = \frac{7 - 11}{-2} = 2,$$

ezért

$$P(x) = -x^2 + 8x - 4 + 2 \cdot (x - 1)(x - 2)(x - 4) = -x^2 + 8x - 4 + 2 \cdot (x^3 - 7x^2 + 14x - 8) = 2x^3 - 15x^2 + 36x - 20$$

1.3. Hermite interpoláció

Most olyan polinomot keresünk, amelynek az x_1, x_2, \dots, x_n helyeken nem csak az értéke adott, hanem a deriváltja is. Tehát keresünk olyan $H(x)$ polinomot, amely $i = 1, 2, \dots, n$ esetén $H(x_i) = y_i$ és $H'(x_i) = y'_i$.

1.4. Tétel.

Ha x_1, x_2, \dots, x_n különböző valós számok, akkor van olyan legfeljebb $2n-1$ -ed fokú polinom, amelynek értéke az x_i helyen éppen y_i és deriváltja az x_i helyen y'_i , vagyis $H(x_i) = y_i$ és $H'(x_i) = y'_i$ minden $i = 1, 2, \dots, n$, esetén.

1.5. Tétel.

A fenti tételnek eleget tevő polinom egyértelmű.

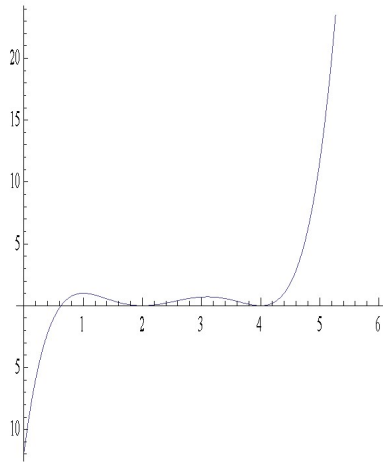
Az Hermit-féle interpolációs polinomot - hasonlóan a Lagrange interpolációhoz - alappolinomok segítségével állítjuk elő. Kétféle alappolinomot használunk.

A $h_{i,0}(x)$ alappolinom az x_i hely kivételével 0 értéket vesz fel, az x_i helyen 1 az értéke, a deriváltja pedig minden alappontban 0.

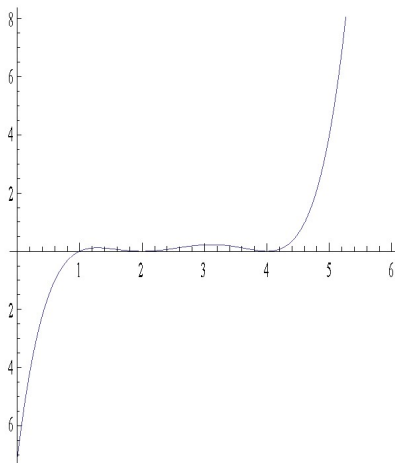
A $h_{i,1}(x)$ alappolinom értéke minden alappontban 0, deriváltja pedig az x_i hely kivételével 0 értéket vesz fel, az x_i helyen pedig 1 az értéke. Tehát:

$$\begin{aligned} h_{i,0}(x) &= \begin{cases} 0 & \text{ha } x = x_j \quad j = 1, 2, \dots, i-1, i+1, \dots, n \\ 1 & \text{ha } x = x_i \end{cases} \\ h'_{i,0}(x_i) &= 0 \quad \text{ha } i = 1, 2, \dots, n. \end{aligned} \quad (4)$$

$$\begin{aligned} h_{i,1}(x_i) &= 0 \quad \text{ha } i = 1, 2, \dots, n. \\ h'_{i,1}(x) &= \begin{cases} 0 & \text{ha } x = x_j \quad j = 1, 2, \dots, i-1, i+1, \dots, n \\ 1 & \text{ha } x = x_i \end{cases} \end{aligned} \quad (5)$$



6. ábra. A $h_{1,0}(x)$ Hermit alappolinom az $x = 1, 2, 4$ alappontokra



7. ábra. A $h_{1,1}(x)$ Hermit alappolinom az $x = 1, 2, 4$ alappontokra

1.6. Tétel.

Az Hermit-féle alappolinomok segítségével az interpolációs polinom:

$$H(x) = \sum_{i=1}^n y_i \cdot h_{i,0}(x) + \sum_{i=1}^n y'_i \cdot h_{i,1}(x)$$

Az Hermit-féle interpolációs polinomokat a Lagrange-féle interpolációs polinomok segítségével állítjuk elő.

1.7. Tétel.

$$h_{i,0}(x) = [l_i(x)]^2 \cdot \left[1 - \frac{\omega''(x_i)}{\omega'(x_i)} \cdot (x - x_i) \right]$$

és

$$h_{i,1}(x) = [l_i(x)]^2 \cdot (x - x_i).$$

1.4. Megjegyzés.

Az Hermit-féle interpoláció általánosítható úgy, hogy az x_i alappontokban nem csak a függvényértéket és az első derivált értékét írjuk elő, hanem az első k derivált értékét. Ekkor egy legfeljebb $(k+1) \cdot n - 1$ -edfokú interpolációs polinom állítható elő.

Abban a speciális esetben, ha $n = 1$ és $k > 1$, a Taylor polinomhoz jutunk:

$$T_k(x) = y_1 + y_1' \cdot (x - x_1) + y_1'' \cdot \frac{(x - x_1)^2}{2!} + \dots + y_1^{(k)} \cdot \frac{(x - x_1)^k}{k!}.$$

1.5. Példa.

Illesszünk interpolációs polinomot az $A(1; 6)$, $B(2; 11)$ és $C(4; 27)$ pontokra úgy, hogy a derivált az A pontban 136, a B pontban -102 és a C pontban 490 legyen!

Megoldás: Most egy legfeljebb ötödfokú polinomot keresünk. Előállítjuk a (1.7) tétel szerinti alappolinomokat. Az (1.1) és (1.2) példákból ismerjük már a megfelelő Lagrange alappolinomokat. Felhasználjuk, hogy

$$\begin{aligned}\omega(x) &= (x-1)(x-2)(x-4) \\ \omega'(x) &= (x-2)(x-4) + (x-1)(x-4) + (x-1)(x-2) \\ \omega''(x) &= 2(x-1) + 2(x-2) + 2(x-4)\end{aligned}\tag{6}$$

Ezért

$$\begin{aligned}\omega'(1) &= 3, & \omega'(2) &= -2, & \omega'(4) &= 6, \\ \omega''(1) &= -8, & \omega''(2) &= -2, & \omega''(4) &= 10\end{aligned}$$

Tehát

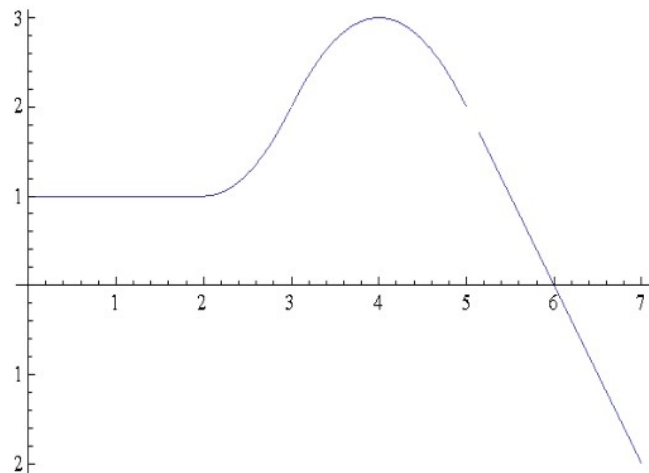
$$\begin{aligned}h_{1,0}(x) &= \left[\frac{x^2-6x+8}{3} \right]^2 \cdot \left[1 + \frac{8}{3}(x-1) \right] \\ h_{2,0}(x) &= \left[\frac{x^2-5x+4}{-2} \right]^2 \cdot [1 - (x-2)] \\ h_{3,0}(x) &= \left[\frac{x^2-3x+2}{6} \right]^2 \cdot \left[1 - \frac{10}{6}(x-4) \right]\end{aligned}\tag{7}$$

Továbbá

$$\begin{aligned}h_{1,1}(x) &= \left[\frac{x^2-6x+8}{3} \right]^2 \cdot (x-1) \\ h_{2,1}(x) &= \left[\frac{x^2-5x+4}{-2} \right]^2 \cdot (x-2) \\ h_{3,1}(x) &= \left[\frac{x^2-3x+2}{6} \right]^2 \cdot (x-4)\end{aligned}\tag{8}$$

A keresett interpolációs polinom:

$$\begin{aligned}H(x) &= 6 \cdot h_{1,0}(x) + 11 \cdot h_{2,0}(x) + 27 \cdot h_{3,0}(x) + 136 \cdot h_{1,1}(x) - 102 \cdot h_{2,1}(x) + 490 \cdot h_{3,1}(x) = \\ &= x^5 + x^3 - 161x^2 + 450x - 285\end{aligned}\tag{9}$$



8. ábra. Spline függvény. Csatlakozási pontok: $x = 2, 3, 5$.

1.4. Spline

A klasszikus Lagrange-féle interpolációt nem előnyös magas fokszámú polinomok esetén alkalmazni erősen oszcilláló tulajdonsága miatt. A gyakorlatban többnyire sima, kevésbé oszcilláló interpolációs polinomra van szükség. Segíthetne a Hermit-féle interpoláció, ha tudnánk az adott pontokban a deriváltakat is, de ezek a legtöbb esetben ismeretlenek.

Előnyösebb szakaszonként alacsonyabb fokszámú polinomokkal próbálkozni, amelyek megfelelően csatlakoznak. A legegyszerűbb eset, ha az adott pontok közé lineáris függvényeket illesztünk. Ez az Euler-féle töröttvonalak módszere sok esetben segíthet, de más esetekben szükséges, hogy a csatlakozási pontokban ne csak folytonos, hanem differenciálható is legyen a függvény.

Napjainkban legelterjedtebb a "kübös spline interpoláció", amikor is az adott pontok közé harmadfokú polinomokat illesztünk úgy, hogy a csatlakozási pontokban a függvénynek jobbról és balról megegyeznek a deriváltjai és második deriváltjai is.

1.2. Definíció.

Harmadfokú (kübös) interpolációs spline a következő:

Adottak az $a = x_0 < x_1 < \dots < x_n$, b , pontok, ahol a függvényértékek legyenek $f(x_i) = y_i \quad i = 1, \dots, n$.

Az $f(x)$ függvényt interpoláló kübös $S(x)$ spline-t a következő módon definiáljuk:

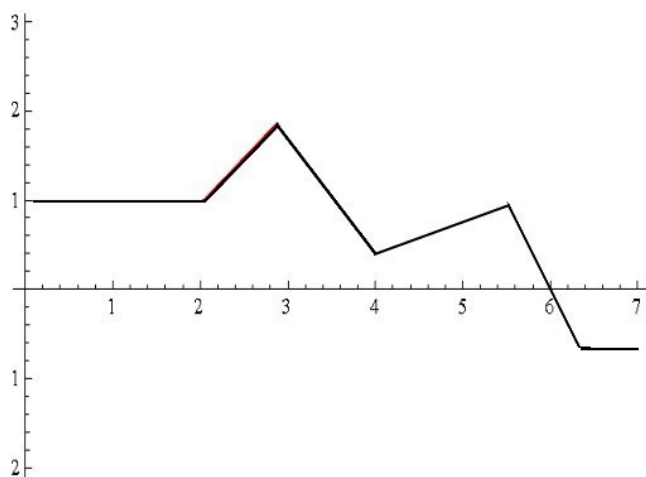
- $S(x)$ szakaszonként harmadfokú polinom úgy, hogy $S_i(x)$ harmadfokú polinom az $[x_i; x_{i+1}] \quad i = 1, \dots, n - 1$ intervallumon.

- $S(x_i) = y_i$ ha $i = 1, \dots, n$ (Tehát $S(x)$ interpolálja az $f(x)$ függvényt)
- $S_i(x_{i+1}) = S_{i+1}(x_{i+1})$ $i = 1, \dots, n - 2$ (Folytonosan csatlakoznak.)
- $S'_i(x_{i+1}) = S'_{i+1}(x_{i+1})$ $i = 1, \dots, n - 2$ ($S(x)$ -nek nincsenek sarkai.)
- $S''_i(x_{i+1}) = S''_{i+1}(x_{i+1})$ $i = 1, \dots, n - 2$ (Görbületek egyenlősége.)
- Továbbá az $(a; x_1)$ és $(x_n; b)$ intervallumokon a függvény elsőfokú.

A spline megkonstruálására többféle módszer ismeretes. Itt most azt a gondolatot folytatjuk, amit a Lagrange és Hermit interpolációnál már láttuk, hogy hatékonyan működik, azaz az interpoláló függvényt alappfüggvények lineáris kombinációjaként állítjuk elő.

1.4.1. Lineáris spline "sátorfüggvényekből"

A feladat most az, hogy az $x_1 < x_2 < \dots < x_n$, pontokban a függvényértékek legyenek y_1, y_2, \dots, y_n . Az adott pontok között a spline függvény elsőfokú, és $x < x_1$ valamint $x > x_n$ esetén a spline függvény legyen konstans.



9. ábra. Spline függvény lineáris szakaszokból.

Definiáljuk az alappfüggvényeket a következő képpen:

1.3. Definíció.

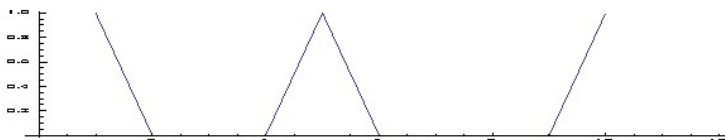
"Sátorfüggvények" (tent functions)

$$l_0(x) = \begin{cases} \frac{x-x_1}{a-x_1} & \text{ha } a \leq x \leq x_1 \\ 0 & \text{egyébként} \end{cases} \quad (10)$$

$j = 1, 2, \dots, n$ esetén

$$l_j(x_i) = \begin{cases} \frac{x-x_{j-1}}{x_j-x_{j-1}} & \text{ha } x_{j-1} \leq x \leq x_j \\ \frac{x-x_{j+1}}{x_j-x_{j+1}} & \text{ha } x_j \leq x \leq x_{j+1} \\ 0 & \text{kívül.} \end{cases} \quad (11)$$

$$l_{n+1}(x) = \begin{cases} \frac{x-x_n}{b-x_n} & \text{ha } x_n \leq x \leq b \\ 0 & \text{egyébként} \end{cases} \quad (12)$$



10. ábra. Sátor függvények

A spline függvény:

$$s(x) = \sum_{i=0}^{n+1} c_i \cdot l_i(x),$$

ahol $c_i = y_i$ ha $i = 1, 2, \dots, n$, valamint $c_0 = -y_1$ és $c_{n+1} = -y_n$.

1.4.2. Lineáris spline "csonkolt" függvényekből

Egy másik lehetséges módszer a spline előállítására az úgynevezett "csonkolt függvények" használata.

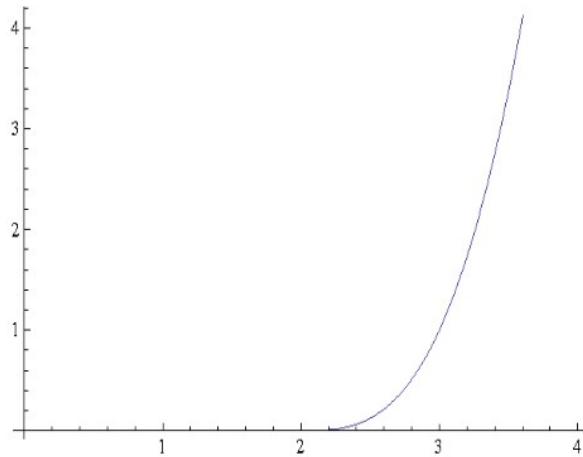
1.4. Definíció.

Legyen $f(x)$ tetszőleges függvény. Csonkolással kapjuk a következő függvényt:

$$f_+(x) = \begin{cases} f(x) & \text{ha } x \geq 0 \\ 0 & \text{ha } x < 0 \end{cases} \quad (13)$$

A spline előállítása:

$$s(x) = c_0 + \sum_{i=1}^n c_i f_+(x - x_i) \quad (14)$$



11. ábra. Az $(x - 2)^3$ függvény csonkolásából kapott függvény

Ebben az egyenletben $n + 1$ darab ismeretlen együttható van. Az $s(x_i) = y_i$ éppen n darab feltétel, továbbá még egy feltételt jelent, hogy az $[x_n; b]$ intervallumon $s(x)$ konstans. Tehát

$$\begin{aligned}
 s(x_1) &= c_0 = y_1 \\
 s(x_2) &= c_0 + c_1 f_+(x_2 - x_1) = y_2 \\
 &\vdots \\
 s(x_n) &= c_0 + \sum_{i=1}^n c_i f_+(x_n - x_i) = y_n \\
 c_1 + c_2 + \dots + c_n &= 0
 \end{aligned} \tag{15}$$

1.4.3. Kúbos spline "csonkolt" függvényekből

A spline előállítását most az $(x - x_i)_+^3$ alapfüggvényekkel történik:

$$s(x) = b_0 + b_1 x + \sum_{i=1}^n c_i (x - x_i)_+^3 \tag{16}$$

alakban. Ebben az egyenletben $n + 2$ darab ismeretlen együttható van. Az $s(x_i) = y_i$ éppen n darab feltétel, továbbá még két feltételt jelent, hogy az $[x_n; b]$ intervallumon $s(x)$ elsőfokú. Tehát

$$\begin{aligned}
 s(x_1) &= b_0 + b_1 x_1 = y_1 \\
 s(x_2) &= b_0 + b_1 x_2 + c_1 (x_2 - x_1)^3 = y_2 \\
 &\vdots \\
 s(x_n) &= b_0 + \sum_{i=1}^n c_i (x_n - x_i)^3 = y_n \\
 c_1 + c_2 + \dots + c_n &= 0 \\
 c_1 x_1 + c_2 x_2 + \dots + c_n x_n &= 0
 \end{aligned} \tag{17}$$

Az utolsó előtti egyenlet abból a feltételből származik, hogy $x > x_n$ esetén a harmadfokú tag együttthatója zérus. Az utolsó egyenlet pedig abból a feltételből származik, hogy $x > x_n$ esetén a másodfokú tag együttthatója is zérus.

1.6. Példa.

Illesszünk köbös spline-t a $(0; 0)$, $(1; 1)$, $(2; 4)$ pontokra.

Megoldás:

$$s(x) = b_0 + b_1x + c_1x_+^3 + c_2(x-1)_+^3 + c_3(x-2)_+^3$$

A feltételek szerint:

$$\begin{aligned} s(0) &= b_0 = 0 \\ s(1) &= b_0 + b_1 + c_1 = 1 \\ s(2) &= b_0 + 2b_1 + 8c_1 + c_2 = 4 \\ c_1 + c_2 + c_3 &= 0 \\ c_2 + 2c_3 &= 0 \end{aligned}$$

Ennek az egyenletrendszernek a megoldása:

$$b_0 = 0, b_1 = \frac{1}{2}, c_1 = \frac{1}{2}, c_2 = -1, c_3 = \frac{1}{2}.$$

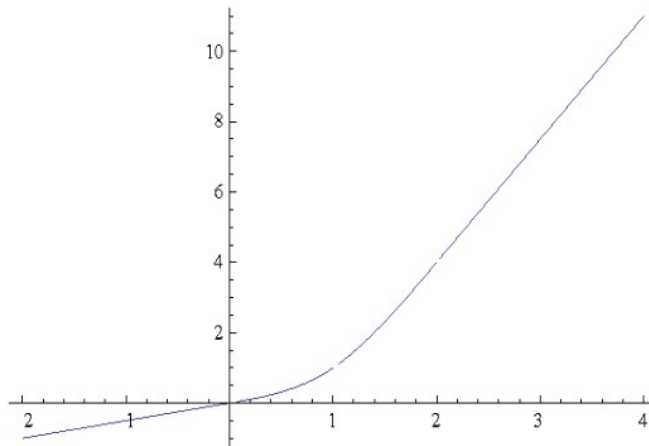
Tehát a spline:

$$s(x) = \frac{1}{2}x + \frac{1}{2}x_+^3 - (x-1)_+^3 + \frac{1}{2}(x-2)_+^3$$

Szakaszonként:

$$\begin{aligned} &\frac{1}{2}x && \text{ha } x < 0 \\ &\frac{1}{2}x^3 + \frac{1}{2}x && \text{ha } 0 \leq x < 1 \\ -\frac{1}{2}x^3 + 3x^2 - \frac{5}{2}x + 1 && \text{ha } 1 \leq x < 2 \\ &\frac{7}{2}x - 3 && \text{ha } 2 \leq x \end{aligned} \tag{18}$$

Könnyen ellenőrizhető, hogy a csomópontokban a jobb és baloldali deriváltak és második deriváltak is megegyeznek.



12. ábra. Az $s(x)$ függvény grafikonja

1.5. Felületek illesztése

Ebben a szakaszban egy kis kitekintés olvasható a felületek interpolációjáról. Forrás: internet.

TÉRBELI INTERPOLÁCIÓ Szerkesztette: Nigel M. Waters, University of Calgary Magyar változat: Závoti József, Geodéziai és Geofizikai Kutató Intézet, Sopron

A. BEVEZETÉS

- térbeli interpoláció az az eljárás, amely a rendelkezésre álló megfigyelések által meghatározott térség mintavétellel nem rendelkező pontjaiban becslést ad a vizsgált tulajdonságok értékére
- a legtöbb esetben a tulajdonságot jellemző értéknek egy meghatározott intervallumba kell esnie, vagy egy meghatározott értékkel vett hányadosát kell tekinteni
- úgy lehet elképzelni, mint a fordítottját annak az eljárásnak, amely egy digitális magasságmodellből kiválasztja azt a néhány pontot, amely megfelelő pontossággal határozza meg a felületet
- a térbeli interpoláció azon a feltevésen alapul, hogy a térben egymáshoz közel elhelyezkedő pontok értéke nagyobb valószínűséggel hasonló, mint az egymástól messze levő pontoké (Tobler törvénye a geográfiára)
- a térbeli interpoláció nagyon fontos alkotórésze sok GIS-nek
- a térbeli interpoláció a GIS-ekben a következő célokra használható:
 - szintvonalak szolgáltatása adatok grafikus megjelenítéséhez
 - a felület valamely jellemzőjének kiszámítása egy adott pontban
 - az összehasonlításához használt mértékegység megváltoztatása, amikor különböző rétegekben különböző adatstruktúrák kerülnek alkalmazásra
 - gyakran használatos segédeszköz térbeli döntéshozatali folyamatoknál fizikai és humán geográfiában egyaránt, továbbá olyan rokon területeken, mint ásványi nyersanyagok feltárása és szénhidrogénkutatás
- a térbeli interpoláció sok eljárása az eredetileg idősorok analízisére kifejlesztett egydimenziós

módszerek kétdimenzióra történt továbbfejlesztése

- a jelen fejezet a térbeli interpoláció bevezetését tartalmazza és megvizsgálja a pontokra alapozott interpolációt, míg a következő a tartományi eljárásokkal és néhány alkalmazással foglalkozik

B. AZ INTERPOLÁCIÓS ELJÁRÁSOK OSZTÁLYOZÁSA

- a térbeli interpolációs eljárásokat különböző módon osztályozhatjuk:

1. Pontok interpolációja/Tartományi interpoláció

- pontokra alapozott

- adva van bizonyos számú pont, amelyeknek helye és értéke ismert; meghatározandók bizonyos előre megadott helyen lévő más pontok értékei

- pont interpolációt olyan adatoknál alkalmaznak, amelyek pontszerű helyeken gyűjthetők, pl. meteorológiai állomások által szolgáltatott adatoknál, földrajzi pontok magasságainál, olajkutaknál mért adatoknál, porozitási méréseknél

- az ilyen interpolációval kapott rácspontok gyakran szolgálnak inputként számítógépes szintvonalrajzoló algoritmusokhoz

- a rácspontok meghatározása után az izovonalak (pl. szintvonalak) már átvezethetők közöttük oly módon, hogy minden rácspont-pár által meghatározott egyenesszakaszon lineáris interpolációt alkalmazunk

- a pontról pontra végzett interpoláció a GIS-ben a leggyakrabban végzett térbeli interpoláció típus

- vonalokról pontokra

- pl. szintvonalokról magassági rácsra

- adott tartományú interpoláció

- forrás-zónák egy bizonyos halmazán térképezett adatok egy adott halmaza esetén meghatározandók az adatok értékei forrás-zónák egy másik halmazára

- pl. népszámlálási kerületekben kapott népesség-számokból megbecsülendők a választási kerületek népességszámjai

2. Globális/lokális interpolátorok

- a globális interpolátorok egyetlen függvényt határoznak meg, amely az egész területre vonatkozik

- egy input érték megváltoztatása az egész térképre hatással van

- a lokális interpolátorok ismételten alkalmaznak egy algoritmust a teljes ponthalmaznak egy kis részére

- egy input érték megváltoztatása csak az ablakon belüli eredményre van hatással

- a globális algoritmusok általában simább felületeket eredményeznek, amelyeken a változások kevésbé hirtelen jellegűek

- akkor alkalmazzák, ha rendelkezésre áll egy hipotézis a felület alakjáról, pl. egy trend esetén

- vannak olyan lokális interpolátorok, amelyek kiterjeszthetők úgy, hogy magukba foglalják az adott pontok egy nagy hányadát, ami által bizonyos értelemben globálissá válnak

- ezért nincs éles határ a globális és a lokális interpolátorok között

- ez félreértésekhez és vitákhoz vezetett az irodalomban

3. Egzakt/közelítő interpolátorok

- az egzakt interpolátorok az interpoláció alapját képező adatpontokon az eredeti értékeket hűen (eltérés nélkül) adják vissza

- a felület áthalad mindazon pontokon, amelyek értéke ismert
- az adatpontokon való hű értékvi visszaadás fontos követelmény sok alkalmazásnál, pl. az olajiparban
- a közelítő interpolátorok, a B-spline-ok és a Krigelés mind olyan módszerek, amelyek hű értékvi visszaadást biztosítanak az adott pontokban
- a Krigelésnél, mint alább látni fogjuk, ún. nugget-effektus léphet fel, ilyen esetben az egzakt interpolátor megjelölés már nem megfelelő
- a közelítő interpolátorokat olyan esetekben alkalmazzák, amikor az adott felületi értékek bizonyos mértékben bizonytalanok
- itt az a nézet kerül alkalmazásra, hogy sok adathalmaz esetében léteznek lassan változó globális trendek, és ezekhez a trendekhez lokális fluktuációk adódnak, melyek viszont gyors változásúak, és így bizonytalanságot (hibát) eredményeznek a rögzített értékekben
- a simítás ezért csökkenti a hibák hatását az eredő felületre

4. Sztohasztikus/determinisztikus interpolátorok

- a sztohasztikus módszerekben benne van a véletlen fogalma
- az interpolált felületet úgy lehet elképzelni, mint egyet azok közül, amelyek az ismert adatpontokra szorítkozva megfigyelhetők lettek volna
- a sztohasztikus módszerek alkalmazzák a trend-felület analízist, a Fourier analízist és a Krigelést
- a trend-felület analízishez hasonló módszerek lehetővé teszik a felület statisztikai szignifikanciájának és az előrejelzett értékek bizonytalanságának a kiszámítását
- a determinisztikus módszerek nem alkalmaznak valószínűségelméletet (pl. a közelítőek)

5. Fokozatos/hirtelen változású interpolátorok

- fokozatos interpolátoroknak tipikus példája a távolsággal súlyozott mozgó átlag
- rendszerint fokozatos változású interpolált felületet eredményez
- azonban, ha a mozgó átlag módszerénél a használt pontok száma erősen lecsökken, vagy akár egy lesz, akkor a felületben hirtelen változások lesznek
- szükségessé válhat akadályok beiktatása az interpolációs folyamatba
- félig áteresztő, pl. időjárási frontok
- gyorsan változó, de folytonos értékeket fognak eredményezni
- áthatolhatatlan akadályok, pl. geológiai törésvonalak
- hirtelen változásokat fognak eredményezni

C. PONTOKRA ALAPOZOTT INTERPOLÁCIÓ - EGZAKT MÓDSZEREK

- Lam (1983) és Burrough (1986) ismertettek egy sor kvantitatív interpolációs módszert, amelyek alkalmazhatók számítógépes szintvonalrajzoló algoritmusokban
- ebben és a következő részben ezeket osztályozni fogjuk egzakt és közelítő módszerekre
- a jelen rész az egzakt módszerekkel foglalkozik

1. Közelítő (proximális)

- minden érték egyenlőnek tekintendő a legközelebbi ismert ponttal
- ez egy lokális interpolátor
- aránylag kevés számítás igényel
- az output adatstruktúra Thiessen poligonokból áll, a határoknál hirtelen változásokkal
- ökológiai alkalmazásai vannak, mint pl. területek és hatás-zónák
- nominális adatokra a legjobb, bár Thiessen eredetileg arra használta, hogy eső-csapadék ada-

tokból tartományi becsléseket számítsion

- abszolút robusztus, mindig eredményt szolgáltat, de nincs semmi "intelligenciája" a vizsgált rendszerre vonatkozóan
- nagyon kevés térképészeti programcsomag tartalmazza, egy említésre méltó kivétel a SYMAP

2. B-spline-ok

- szakaszonkénti polinomokat használnak oly módon, hogy egymáshoz folytonos első és második deriváltakkal csatlakozó felületdarabkákat állítsanak elő
- folytonos eredményt adnak a következő értékekre:
 - magasság (0-adrendű folytonosság) - A felület nem tartalmaz sziklákat
 - meredekség (elsőrendű folytonosság) - A meredekség nem változik hirtelen, a szintvonalak simák
 - görbület (másodrendű folytonosság) - Minimális görbületet szolgáltat
 - minimális görbületű folytonos felületet szolgáltatnak
 - az output adatstruktúra raszteren levő pontokból áll
 - vegyük észre, hogy a maximumok és a minimumok nem feltétlenül az adatpontokon lesznek
 - lokális interpolátor
 - lehet egzakt, vagy használható felületsimításra
 - számítási igénye közepes
 - nagyon sima felületekre a legjobbak
 - erős fluktuációjú felületeken nem jók, mert a spline-okban vad oszcillációt okozhatnak
 - népszerűek az általános felület-interpolációs programcsomagokban, de GIS-ben ritkán fordulnak elő
 - közelíthetők olyan szintvonalak simításával, amelyek TIN modellből lettek rajzolva
 - lásd Burrough (1986), Davis (1986), továbbá matematikai vonatkozásokhoz Lam (1983), valamint Hearn és Baker (1986)
 - leírásuk megtalálható még a "numerikus approximáció elmélet"-ben

3. Krigelés

- kifejlesztették: Georges Matheron, mint a "regionalizált változók elmélete", valamint D.G. Krige, mint az interpoláció optimális elmélete a bányászatban
- ennek az eljárásnak az szolgál alapjául, hogy a térben a pontok között a szórás milyen ütemben változik
- ez leolvasható a variogram-ról, amely azt mutatja, hogy hogyan függ a pontok közötti távolságtól a pontokhoz tartozó értékek átlagos eltérése

Variogramok

- D_e (függőleges tengely) felírható, mint $E(z_i - z_j)^2$, tehát mint az eltérés "várható értéke"
- tehát egymástól d távolságra levő bármely két pont magasságértékének átlagos eltérése
- d (vízszintes tengely) értéke az i és j közötti távolság
- a legtöbb variogram az ábrához hasonló viselkedést mutat
- D_e felső határát (aszimptotáját) küszöbnek hívják
- azt a távolságot, amelynél ez a határérték fellép, hatástávolságnak (range) nevezik
- az y -tengellyel való metszés pontot "nugget"-nek hívják
- ha a nugget nullától különbözik, az arra utal, hogy ugyanazon a ponton végzett ismételt méré-

sek különböző értékeket szolgáltatnak

- a variogram készítésekor szükséges néhány feltevést tenni a felületen megfigyelt változás jellegét illetően:

- az egyszerű Krigelés feltételezi, hogy a felület középértéke állandó, trend nem lép fel és minden változás statisztikai

- az univerzális Krigelés feltételezi, hogy a statisztikai változáson kívül a felület egy determinisztikus trend-el is rendelkezik

- akármelyik esetben, ha a trendek már meg lettek állapítva (vagy a feltételezés szerint nincs trend), minden további változásnál feltesszük, hogy a távolság függvényei

A variogram származtatása

- a krigelés input adatait rendszerint szabálytalanul elhelyezkedő mintapontok alkotják

- egy variogram kiszámításához meg kell határoznunk, hogy a szórás hogyan növekszik a távolság függvényében

- először is a hatástávolságot diszkrét intervallumokra osztjuk, pl. a vizsgált területen 0 és a maximális távolság között 10 intervallumot veszünk fel

- minden pontjára kiszámítjuk a távolságot és a z értékek különbségének a négyzetét

- minden párt hozzárendelünk a hatástávolságok egyikéhez, és kiszámítjuk a teljes szórást mind-egyik hatástávolságra

- a hatástávolság középpontjában kirajzoltatjuk a kapott értéket A becslések számítása

- miután elkészült a variogram, az interpoláció távolság-súlyainak becslésére használjuk fel

- az interpolált értékek egyenlők néhány ismert pont súlyozott értékeinek összegeivel, ahol a súlyok az interpolált és az ismert pontok közötti távolság függvényei

- a súlyokat úgy választjuk meg, hogy a becslések

- torzítatlanok (ismételt alkalmazás esetén a Krigelés átlagban helyes eredményt adna)

- minimális szórásúak (az ismételt becslések közötti változás minimális)

- az eljárással kapcsolatos problémák:

- nagyszámú adatpont esetén ez az eljárás nagyon számítás-igényes

- a variogram becslése nem egyszerű, egyetlen eljárást sem lehet a legjobbnak tekinteni

- mivel a változás statisztikai jellegéről több lényeges feltevést kell tenni, az így kapott eredmények nem tekinthetők abszolútnak

- az egyszerű Krigelésre rutinokat tartalmaznak a következők: a Surface II programcsomag (Kansas Geological Survey) és Surfer (Golden Software), továbbá a GEOEAS programcsomag, amelyet PC-re fejlesztettek ki az Amerikai Környezetvédelmi Hivatalban (US Environmental Protection Agency), GENSTAT (.N.AG /The Numerical Algorithms Group LTD/ Anglia)

4. Kézi ("ránézéses") interpoláció

- geográfusok és térképészek hagyományosan nem becsülik sokra

- azonban Dutton-Marion (1988) rámutatott, hogy a geológusok körében ez nagyon fontos eljárás, és hogy a legtöbb geológus valójában bizalmatlan a bonyolultabb, többet tudó algoritmusokkal szemben

- úgy érzik, hogy támaszkodhatnak a szakértelmükre, modellezési képességükre és tapasztalataikra, és ennek a tudásnak a birtokában a geológiai felület előállításába valóságosabb interpolációt integrálhatnak

- történnék kísérletek abba az irányba, hogy ismerettervező eljárásokkal ezt a tudást a szakértőtől összegyűjtsék, és interpoláció céljára szakértő rendszerekbe építsék be

- I. a 74-ik fejezetet, ahol erről a témáról több található
- ez a módszer a következő jellemzőkkel rendelkezik:
- ezek az eljárások lokálisak, ugyanis a szakértő a térkép különböző részein különböző módszereket használhat
- általában az adatpontokon hű értékvisszaadást biztosítanak
- ezekkel a módszerekkel könnyebben modellezhetők olyan hirtelen változások, mint pl. a törszonalak
- a felületek szubjektívek, másik szakértő esetén a felület is más lesz
- az output adatstruktúra általában szintvonal képében jelentkezik

D. PONTOKRA ALAPOZOTT INTERPOLÁCIÓ - KÖZELÍTŐ MÓDSZEREK

1. Trend-felület analízis

- a felületet polinommal közelítjük
- az output adatstruktúra egy polinom-függvény, amely felhasználható raszteren levő rácspontok értékeinek becslésére vagy bármely helyhez tartozó érték becslésére
- a felület tetszőleges (x,y) pontjának z magassága egy x és y hatványait tartalmazó képlettel számítható
- például egy dőlt sík felületet egy lineáris (elsőfokú) egyenlet határoz meg:

$$z = a + bx + cy$$

- pl. egy másodfokú egyenlet egy egyszerű dombot vagy völgyet ír le:

$$z = a + bx + cy + dx^2 + exy + fy^2$$

- általános érvényű, hogy egy n-edfokú felület bármely keresztmetszetének legfeljebb n-1 váltakozó maximuma és minimuma lehet
- pl. egy harmadfokú felület bármely keresztmetszetének egy maximuma és egy minimuma lehet
- a harmadfokú felület egyenlete:

$$z = a + bx + cy + dx^2 + exy + fy^2 + gx^3 + hx^2y + ixy^2 + jy^3$$

- a trend-felület egy globális interpolátor
- feltételezi, hogy a felület általános trendje független az egyes mintavételi pontokban fellépő véletlen hibáktól
- számításigénye aránylag kicsi
- problémák
- a modelltől tett statisztikai feltevések a gyakorlatban ritkán teljesülnek
- a tartományok szélein igen komoly problémák léphetnek fel
- a polinomiális modell lekerekített felületet eredményez
- humán és fizikai alkalmazások esetén ez ritkán felel meg a valóságnak
- igen sok térképészeti programcsomag tartalmazza
- nem ortogonális polinomokra l. Davis (1973) és Sampson (1978): ortogonális polinomokra pedig Mather (1976)

2. Fourier-sorok

- a felületet sinus és cosinus hullámok egy sorának összetevésével (overlay) közelíti
- globális interpolátor
- számításiigénye közepes
- az output adatstruktúrát a Fourier-sor képezi, amely felhasználható raszteres rácspontról értékek, vagy tetszőleges pontbeli érték becslésére
- a legjobban olyan adathalmazok esetén használható, amelyek jellegzetes periodicitást mutatnak, pl. a tenger hullámai
- ritkán található meg számítógépes programcsomagokban
- egyszerű programot és leírást közöl Davis (1973)

3. Változó átlag/távolsággal súlyozott átlag

- a becslések n ismert pontbeli értékek átlagaival egyenlők:

$$z = \sum w_i z_i / \sum w_i$$

ahol w a távolság valamilyen függvénye, pl.:

$$w = 1/dkw = e^{-kd}$$

- szinte végtelen sokfajta algoritmus használható, az eltéréseket többek között az alábbiak jelentik:
- a távolságfüggvény jellege
- a felhasznált pontok számának változtatása
- az az irány, amelyből a pontok ki lesznek választva
- a legelterjedtebben használt módszer
- az ezzel a módszerrel szembeni ellenvetések abból a tényből erednek, hogy az interpolált értékek az adatok által meghatározott intervallumba kell, hogy essenek
- egyetlen interpolált érték sem eshet a megfigyelt z értékek intervallumán kívül
- további problémák:
- az átlagoláshoz hány pont legyen felhasználva?
- a szabálytalanul elhelyezkedő pontok esetén mi a teendő?
- a széleken fellépő jelenségek hogyan kezelendők?

IRODALOM

Burrough, P.A., 1986. "Principles of Geographical Information System for land Resources Assessment", Clarendon, Oxford. See. Chapter 8

Davis, J.C., 1986. "Statistics and Data Analysis in Geology", 2nd edition, Wiley, New York. (Also see first, 1973, edition for program listings.)

Dutton-Marion, K.E., 1988. "Principles of Interpolation procedures in the Display and Analysis of Spatial Data: A Comparative Analysis of Conceptual and Computer Contouring", unpublished Ph.D. Thesis, Department of Geography, University of Calgary, Calgary, Alberta

Hearn, D., and Baker, M.P., 1986. "Computer Graphics", Prentice-Hall Inc, Englewood Cliffs, N.J
Jones, T.A., Hamilton, D.E. and Johnson, C.R., 1986. "Contouring Geologic Surfaces with the Computer", Van Nostrand Reinhold, New York

- Lam, N., 1983. "Spatial Interpolation Methods: A Review", *The American Cartographer* 10(2):129-149
- Mather, P.M., 1976. "Computational Methods of Multivariate Analysis in Physical Geography", Wiley, New York
- Sampson, R.J., 1978. "Surface" II, revised edition, Kansas Geological Survey, Lawrence, Kansas
- Waters, N.M., 1988. "Expert System and Systems of Experts", Chapter 12 in: W.J. Coffey, ed., "Geographical Systems and Systems of Geography: Essays in Honour of William Warntz", Department of Geography, University of Western Ontario, London, Ontario

2. Végtelen sorok

Végtelen sok tagú összegek tulajdonságait vizsgáljuk. Először végtelen sok szám összegét, majd végtelen sok függvény összegét. Kulcskérdés, hogy egy végtelen sok tagú összeg milyen feltételek mellett viselkedik úgy, mint egy véges sok tagú összeg.

2.1. Numerikus sorok

2.1. Definíció.

Az $a_1, a_2, \dots, a_n, \dots$ számokból alkotott

$$a_1 + a_2 + \dots + a_n + \dots = \sum_{i=1}^{\infty} a_i$$

végtelen összeget (szám)sornak nevezzük.

2.2. Definíció.

$$\begin{aligned} s_1 &= a_1 \\ s_2 &= a_1 + a_2 \\ &\vdots \\ s_n &= a_1 + a_2 + \dots + a_n \\ &\vdots \end{aligned} \tag{19}$$

sorozatot a $\sum_{i=1}^{\infty} a_i$ sor részletösszeg sorozatának nevezzük.

2.3. Definíció.

A $\sum_{i=1}^{\infty} a_i$ sor konvergens ha a részletösszeg sorozata konvergens.

Ha konvergens a sor, akkor összege $\lim_{n \rightarrow \infty} s_n = S$.

Ha a részletösszeg sorozat divergens, akkor a sor is divergens.

Először is azt a kérdést vizsgáljuk, hogyan lehet egy sorról megállapítani, hogy konvergens. Másodszor arra keressük a választ, mikor milyen feltételek mellett öröklődnek át a véges összeg azon kellemes tulajdonságai a végtelen sorokra, hogy az összeg tagjai felcserélhetők, tetszőlegesen zárójelezhetők.

2.1. Tétel.

A $\sum_{i=1}^{\infty} a_i$ sor konvergenciájának szükséges feltétele, hogy $\lim_{n \rightarrow \infty} a_n = 0$ legyen.

Bizonyítás:

Ha a sor konvergens, akkor

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} s_n - \lim_{n \rightarrow \infty} s_{n-1} = S - S = 0.$$

Megmutatjuk, hogy a $\lim_{n \rightarrow \infty} a_n = 0$ feltétel nem elégséges.

2.1. Példa.

Példa divergens sorra:

Tekintsük a $\sum_{i=1}^{\infty} \frac{1}{i}$ sort. Itt nyilván teljesül a $\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \frac{1}{n} = 0$ feltétel. Ugyanakkor a sor divergens, mert a részletösszeg sorozata minden határon túl növekedik (végtelenhez tart). Ehhez elég csak azt belátni, hogy a részletösszeg sorozat kettőhatványokhoz tartozó rész-sorozata s_{2^k} végtelenhez tart. Ezt pedig a következőképpen mutatjuk meg:

$$s_{2^k} = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \dots + \frac{1}{8}\right) + \dots + \left(\frac{1}{2^{k-1}+1} + \dots + \frac{1}{2^k}\right)$$

Itt minden zárójelben szereplő rész nagyobb mint $\frac{1}{2}$, mert

$$\frac{1}{3} + \frac{1}{4} > \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$\frac{1}{5} + \dots + \frac{1}{8} > \frac{1}{8} + \dots + \frac{1}{8} = 4 \cdot \frac{1}{8} = \frac{1}{2}$$

és így tovább. Tehát

$$s_{2^k} > 1 + k \cdot \frac{1}{2}. \quad *$$

2.2. Példa.

Példa konvergens sorra.

A

$$\sum_{i=0}^{\infty} q^i = 1 + q + q^2 + \dots + q^i + \dots$$

végtelen (geometriai) sor $|q| < 1$ esetén konvergens. Az n -edik részletösszeg

$$s_n = 1 + q + q^2 + \dots + q^{n-1} = \frac{1 - q^n}{1 - q}$$

Ha $|q| < 1$, akkor $\lim_{n \rightarrow \infty} q^n = 0$, tehát

$$\lim_{n \rightarrow \infty} s_n = \frac{1}{1 - q}$$

Továbbá könnyen látható, hogy $|q| \geq 1$ esetén a sor divergens.

2.4. Definíció.

A (2.3) definíció szerint egy sor akkor és csak akkor konvergens, ha a részletösszeg sorozata konvergens. Sorozatok konvergenciájára vonatkozó Cauchy féle konvergencia kritériumot most megfogalmazzuk a sorokra is.

2.2. Tétel.

Egy $\sum a_i$ sor akkor és csak akkor konvergens, ha minden $\epsilon > 0$ számhoz található N küszöb-szám úgy, hogy bármely $n > N$ és $m > N$ esetén teljesül a következő egyenlőtlenség:

$$|s_m - s_n| = |a_{n+1} + a_{n+2} + \dots + a_m| < \epsilon.$$

Következmény: Ha egy sorhoz hozzáveszünk véges sok tagot, vagy elveszünk belőle véges sok tagot, akkor a konvergens sor konvergens, a divergens sor divergens marad.

2.3. Tétel.

Ha a $\sum_{i=1}^{\infty} |a_i|$ sor konvergens, akkor a $\sum_{i=1}^{\infty} a_i$ sor is konvergens.

Bizonyítás:

Ha a $\sum_{i=1}^{\infty} |a_i|$ sor konvergens, akkor a Cauchy kritérium alapján minden $\epsilon > 0$ számhoz létezik N úgy, hogy $n > N$ és $m > N$ esetén

$$||a_{n+1}| + |a_{n+2}| + \dots + |a_m|| < \epsilon$$

De akkor

$$|a_{n+1} + a_{n+2} + \dots + a_m| < ||a_{n+1}| + |a_{n+2}| + \dots + |a_m|| < \epsilon$$

miatt

$$|a_{n+1} + a_{n+2} + \dots + a_m| < \epsilon$$

is teljesül. Tehát $\sum_{i=1}^{\infty} a_i$ sor is konvergens.

2.5. Definíció.

Ha a $\sum_{i=1}^{\infty} |a_i|$ sor konvergens, akkor a $\sum_{i=1}^{\infty} a_i$ sort abszolút konvergensnek nevezzük.

2.6. Definíció.

Azokat a *konvergens sorokat*, amelyek nem abszolút konvergensnek feltételesen konvergensnek nevezzük.

2.3. Példa.

A (2.2) példában mutatott konvergens sor $|q| < 1$ esetén, egyben abszolút konvergens is.

2.4. Példa.

A következő sor feltételesen konvergens:

$$\sum_{i=1}^{\infty} (-1)^{i+1} \frac{1}{i}$$

A konvergencia a később bizonyítandó (2.11) tételből fog következni. A tagok abszolút értékeiből alkotott sor divergenciáját a (2.1) példában láttuk.

A (2.3) tétel mutatja, hogy a pozitív (nem negatív) tagú sorok különlegesen fontos szerepet játszanak a végtelen sorok között. Ezért a következő tételek ilyen sorokra vonatkoznak.

2.1.1. Pozitív tagú sorok konvergencia-kritériumai

Láttuk, hogy a $\sum a_i$ sor konvergenciájának szükséges feltétele, hogy

$$\lim_{i \rightarrow \infty} a_i = 0$$

legyen. Most elégséges feltételeket fogunk keresni.

Ha a sor tagjai nem negatív számok, azaz minden $a_i \geq 0$, akkor az

$$s_1, s_2, \dots, s_n, \dots$$

részletösszeg sorozat monoton növekedő. Tudjuk, hogy minden monoton és korlátos sorozat konvergens. Elég lesz tehát arra koncentrálni, hogy milyen feltételek biztosítják a részletösszeg sorozat korlátosságát.

2.4. Tétel.

Majoráns kritérium Ha minden $i \in N$ esetén $0 \leq a_i \leq b_i$ teljesül és a $\sum b_i$ sor konvergens, akkor a $\sum a_i$ sor is az.

Bizonyítás:

Ha a $\sum a_i$ sor részletösszeg sorozata $s_1, s_2, \dots, s_n, \dots$ és a $\sum b_i$ sor részletösszeg sorozata pedig $S_1, S_2, \dots, S_n, \dots$, akkor a feltétel szerint

$$\begin{aligned} s_1 &\leq S_1 \\ s_2 &\leq S_2 \\ &\vdots \\ s_n &\leq S_n \\ &\vdots \end{aligned} \tag{20}$$

Mivel az S_n sorozat monoton, korlátos és konvergens, ezért az s_n sorozat is ilyen. *

2.5. Tétel.

Minoráns kritérium Ha minden $i \in N$ esetén $0 \leq b_i \leq a_i$ teljesül és a $\sum b_i$ sor divergens, akkor a $\sum a_i$ sor is az.

Bizonyítás: Most az előző tétellel ellentétben azt tudjuk, hogy

$$\begin{aligned} S_1 &\leq s_1 \\ S_2 &\leq s_2 \\ &\vdots \\ S_n &\leq s_n \\ &\vdots \end{aligned} \tag{21}$$

Mivel a S_n sorozat most minden határon túl növekvő, ezért a s_n sorozat is ilyen. *

2.6. Tétel.

Hányadoskritérium I.

$\alpha)$ Ha minden $i \in N$ esetén $a_i > 0$ és $\frac{a_{i+1}}{a_i} \leq q < 1$, akkor a $\sum a_i$ sor konvergens.

$\beta)$ Ha pedig minden $i \in N$ esetén $\frac{a_{i+1}}{a_i} \geq q > 1$, akkor a $\sum a_i$ sor divergens.

Bizonyítás:

Vizsgáljuk meg a részletösszeg sorozatot. Ha az $\alpha)$ feltétel teljesül, akkor

$$\begin{aligned} a_2 &\leq a_1 q \\ a_3 &\leq a_2 q \leq a_1 q^2 \\ &\vdots \\ a_n &\leq a_{n-1} q \leq a_1 q^{n-1} \\ &\vdots \end{aligned} \tag{22}$$

Tehát a sor egy konvergens geometriai sorral majorálható, ezért konvergens. Ha a $\beta)$ feltétel teljesül, akkor a fordított egyenlőtlenség jelek érvényesek és $q > 1$ miatt a sor divergens geometriai sorral minorálható, azaz maga is divergens. *

2.7. Tétel.

Hányadoskritérium II.

$\alpha)$ Ha minden $i \in N$ esetén $a_i > 0$ és $\lim_{i \rightarrow \infty} \frac{a_{i+1}}{a_i} = r < 1$, akkor a $\sum a_i$ sor konvergens.

$\beta)$ Ha pedig minden $i \in N$ esetén $\lim_{i \rightarrow \infty} \frac{a_{i+1}}{a_i} = r > 1$, akkor a $\sum a_i$ sor divergens.

Bizonyítás:

Ha az $\alpha)$ feltétel teljesül, akkor az

$$\epsilon = \frac{1-r}{2}$$

választáshoz tartozik egy olyan N küszöbindex, amelynél nagyobb i index esetén az $\frac{a_{i+1}}{a_i}$ sorozat tagjai kisebbek, mint $r + \epsilon$, azaz

$$\frac{a_{i+1}}{a_i} < r + \frac{1-r}{2} = \frac{1+r}{2} < 1$$

Következés képpen az első N tag elhagyásával keletkező sor a tétel értelmében konvergens. Ekkor persze az eredeti sor is konvergens, mert véges sok tag elhagyása vagy hozzávétele nem befolyásolja a konvergencia tényét. Ha a β) feltétel teljesül, akkor hasonlóan látható, hogy valamely indextől kezdve teljesül az

$$\frac{a_{i+1}}{a_i} > \frac{1+r}{2} > 1$$

egyenlőtlenség, tehát a sor divergens.

2.1. Megjegyzés.

A hányadoskritérium alapján tehát $\lim \frac{a_{i+1}}{a_i} > 1$ esetben tudjuk, hogy a sor divergens, és $\lim \frac{a_{i+1}}{a_i} < 1$ esetben konvergens a sor. A kritérium nem ad felvilágosítást arról az esetről, amikor a határérték éppen egy. Valóban ekkor az is előfordulhat, hogy a sor divergens és az is, hogy konvergens, amint azt a következő két példa mutatja.

2.5. Példa.

Már láttuk a (2.1) példában), hogy a $\sum_{n=1}^{\infty} \frac{1}{n}$ sor divergens. Ezt azonban a hányadoskritérium nem mutatja, mert

$$\lim_{i \rightarrow \infty} \frac{a_{i+1}}{a_i} = \lim_{i \rightarrow \infty} \frac{i}{i+1} = 1$$

2.6. Példa.

Tekintsük most a következő sort:

$$\sum_{i=1}^{\infty} \frac{1}{i^2}$$

Alkalmazva erre a hányadoskritériumot azt kapjuk, hogy

$$\lim_{i \rightarrow \infty} \frac{a_{i+1}}{a_i} = \lim_{i \rightarrow \infty} \frac{i^2}{(i+1)^2} = 1$$

A kritérium tehát nem mutatja, hogy konvergens-e vagy divergens a sor. Más módszerrel, a majoráns kritérium alkalmazásával megmutatjuk, hogy ez a sor konvergens. Ha $i > 1$, akkor

$$\frac{1}{i^2} < \frac{1}{(i-1)i}$$

Elég tehát megmutatni, hogy a

$$\sum_{i=2}^{\infty} \frac{1}{(i-1)i}$$

sor konvergens. Ennek n -ik részletösszege:

$$\begin{aligned} s_n &= \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \dots + \frac{1}{n(n+1)} = \\ &= 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \dots + \frac{1}{n} - \frac{1}{n+1} = 1 - \frac{1}{n+1} \end{aligned}$$

Ennek van határértéke $n \rightarrow \infty$ esetén. Tehát konvergens.

2.7. Példa.

Az előző két példa azt mutatta, hogy van olyan eset, amikor a hányadoskritérium nem segít. Most arra az esetre mutatunk példát amikor a segítségével könnyen eldönthető, hogy konvergens-e a sor. Tekintsük pl. a következő sort:

2.8. Példa.

$$\sum_{i=1}^{\infty} \frac{i^{10}}{2^i}$$

Ez a sor konvergens, mert

$$\lim_{i \rightarrow \infty} \frac{a_{i+1}}{a_i} = \lim_{i \rightarrow \infty} \left(\frac{i+1}{i} \right)^{10} \cdot \frac{1}{2} = \frac{1}{2}.$$

2.8. Tétel.

Gyökkritérium I.

α) Ha minden $i \in N$ esetén $a_i > 0$ és $\sqrt[i]{a_i} \leq q < 1$, akkor a $\sum a_i$ sor konvergens.

β) Ha pedig minden $i \in N$ esetén $\sqrt[i]{a_i} \geq q > 1$, akkor a $\sum a_i$ sor divergens.

2.9. Tétel.

Gyökkritérium II.

α) Ha minden $i \in N$ esetén $a_i > 0$ és $\lim_{i \rightarrow \infty} \sqrt[i]{a_i} = r < 1$, akkor a $\sum a_i$ sor konvergens.

β) Ha pedig minden $i \in N$ esetén $\lim_{i \rightarrow \infty} \sqrt[i]{a_i} = r > 1$, akkor a $\sum a_i$ sor divergens.

Ez utóbbi két tétel ugyan úgy bizonyítható, mint a hányadoskritérium.

2.10. Tétel.

Integrálkritérium

Legyen az $f(x) > 0$ függvény monoton csökkenő valamely $\alpha \in \mathbf{R}$ valós számtól kezdve és $\lim_{x \rightarrow \infty} f(x) = 0$. A

$$\sum_{n=1}^{\infty} f(n)$$

sor akkor és csak akkor konvergens, ha a

$$\int_{\alpha}^{\infty} f(x) dx$$

improprius integrál is az.

Bizonyítás:

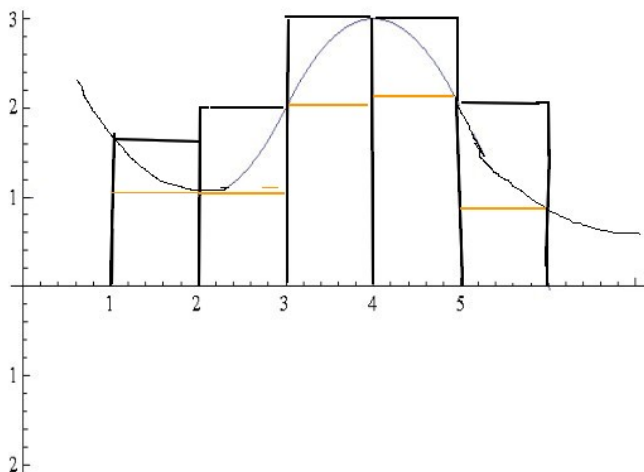
I. eset. Ha az improprius integrál konvergens, akkor a

$$\sum_{n=\beta}^{\infty} f(n)$$

alsó közelítő összege is konvergens, ahol β a legkisebb olyan egész szám, amelyik nagyobb mint α . Ebből következik, hogy a

$$\sum_{n=1}^{\infty} f(n)$$

sor is konvergens.



13. ábra. Integrál-kritérium: Integrál alsó - felső becslése

II. eset. Ha az improprius integrál divergens, akkor a

$$\sum_{n=\beta}^{\infty} f(n)$$

felső közelítő összege is divergens. Ebből következik, hogy a

$$\sum_{n=1}^{\infty} f(n)$$

sor is divergens.

2.1.2. Alternáló sorok

2.7. Definíció.

Ha minden $i \in \mathbb{N}$ esetén $a_i > 0$, akkor az $a_1 - a_2 + a_3 - \dots = \sum (-1)^{i+1} a_i$ sort alternáló (váltakozó előjelű) sornak nevezzük.

2.11. Tétel.

Leibniz tétel

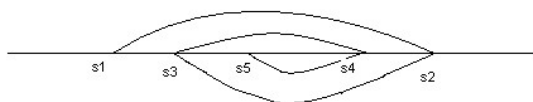
Ha minden $i \in \mathbb{N}$ esetén $a_i > 0$ és az a_i sorozat monoton tart 0-hoz, akkor a $\sum_{i=1}^{\infty} (-1)^{i+1} a_i$ sor konvergens.

Bizonyítás:

A feltételeknek megfelelő sor páratlan indexű részletösszeg sorozata monoton csökkenő, míg a páros indexű rész monoton növekvő. Mindkét rész korlátos, mert minden i -re $s_{2i} < s_1$ és $s_{2i+1} > s_2$. Továbbá

$$\lim_{i \rightarrow \infty} s_{2i+1} - s_{2i} = \lim_{i \rightarrow \infty} a_{2i+1} = 0$$

miatt a részletösszeg sorozat páros és páratlan része ugyan ahhoz a határértékhez tart.



14. ábra. Leibniz típusú sor részletösszegei

2.1.3. Végtelen összegek tulajdonságai

Ebben a szakaszban arra a kérdésre keressük a választ, hogy a véges összeadás mely tulajdonságai, milyen feltételek mellett öröklődnek át a végtelen sorokra. A véges összeg asszociatív, azaz tetszőlegesen zárójellezhető.

2.12. Tétel.

Konvergens sor tetszőlegesen átzárójelezhető, mert az nem változtatja meg az összegét.

Bizonyítás: A zárójelekkel ellátott sor részletösszeg sorozata az eredeti sor részletösszeg sorozatának egy részsorozata. Mivel a feltétel szerint az eredeti konvergens, ezért annak minden végtelen részsorozata is konvergens, és ugyanaz a határértéke. Pl.

$$a_1 + (a_2 + a_3) + a_4 + a_5 + (a_6 + a_7 + a_8) + \dots$$

2.2. Megjegyzés.

Divergens sor megváltozhat a zárójelezéssel. Pl. tekintsük a következő divergens sort:

$$1 - 1 + 1 - 1 + 1 - 1 + 1 - 1 + \dots$$

Ez a sor divergens, mert a részletösszeg sorozata: $1, 0, 1, 0, 1, 0, \dots$. Zárójelezéssel konvergense tehető, pl.:

$$(1 - 1) + (1 - 1) + (1 - 1) + \dots$$

Véges összeg kommutatív, azaz tagjai tetszőleges sorrendbe átrendezhetőek. Ez a tulajdonság végtelen összegre csak akkor öröklődik át, ha a sor abszolút konvergens. Bizonyítás nélkül közöljük a következő két tételt.

2.13. Tétel.

Ha a $\sum a_i$ sor abszolút konvergens és összege S , akkor tetszőlegesen átrendezve is konvergens marad és összege S .

2.14. Tétel.

Ha a $\sum a_i$ sor feltételesen konvergens, akkor átrendezhető úgy, hogy egy előre adott számhoz konvergáljon vagy átrendezhető úgy is, hogy divergens legyen.

2.2. Függvénysorok

Legyen minden $f_n(x)$ függvény értelmezve egy $(a; b)$ intervallumon.

2.8. Definíció.

A $\sum_{n=1}^{\infty} f_n(x)$ összeget *függvénysornak* nevezzük.

A függvénysorok konvergenciáját ugyan úgy definiáljuk, mint a számsorokét.

2.9. Definíció.

A $\sum_{n=1}^{\infty} f_n(x)$ függvénysor konvergens az x helyen, ha itt az $s_n(x) = \sum_{k=1}^n f_k(x)$ részletösszeg sorozat konvergens.

2.10. Definíció.

Azon $x \in (a; b)$ számok halmazát, amelyekre a $\sum_{n=1}^{\infty} f_n(x)$ függvénysor konvergens, konvergencia tartománynak nevezzük. A definíció ekvivalens a következővel:

2.11. Definíció.

A $\sum_{n=1}^{\infty} f_n(x)$ függvénysor konvergens és összege $s(x)$, ha minden $\epsilon > 0$ számhoz található $N_\epsilon(x)$ küszöbindex úgy, hogy $|s(x) - s_n(x)| < \epsilon$, valahányszor $n > N_\epsilon(x)$.

Itt az $N_\epsilon(x)$ jelöléssel azt akartuk kifejezni, hogy a küszöbindex most nem csak ϵ -től, hanem x -től is függhet. Különösen fontos azonban az az eset, amikor egy $(a; b)$ intervallumon van az x változótól független küszöbindex. Egy x -től független küszöbindex akkor található, ha az $N_\epsilon(x)$ korlátos függvénye az x -nek.

2.12. Definíció.

A $\sum_{n=1}^{\infty} f_n(x)$ függvénysort egyenletesen konvergensnek nevezzük az $(a; b)$ intervallumon, ha itt az $N_\epsilon(x)$ korlátos függvénye az x -nek.

Példa egyenletesen és nem egyenletesen konvergens függvénysorra.

2.9. Példa.

Tekintsük a következő függvénysort:

$$\sum_{n=1}^{\infty} \left(\frac{nx}{nx+1} - \frac{(n-1)x}{(n-1)x+1} \right)$$

ahol $x \in [0; 1]$. Ennek a függvénysornak az n -edik részletösszege:

$$s_n(x) = \frac{nx}{nx+1} = 1 - \frac{1}{nx+1}$$

A függvénysor konvergens és összege $s(x)$, ahol

$$s(x) = \lim_{n \rightarrow \infty} s_n(x) = \begin{cases} 0, & \text{ha } x = 0 \\ 1, & \text{ha } x > 0 \end{cases}$$

Rögzítve egy kicsi $\epsilon > 0$ számot

$$|s(x) - s_n(x)| = \begin{cases} 0, & \text{ha } x = 0 \\ \frac{1}{|nx+1|}, & \text{ha } x > 0 \end{cases}$$

Tehát $|s(x) - s_n(x)| < \epsilon$, ha $nx + 1 > \frac{1}{\epsilon}$, azaz

$$n > \frac{\frac{1}{\epsilon} - 1}{x} = N_\epsilon(x)$$

Látható, hogy az $N_\epsilon(x)$ mint x függvénye a $(0; 1]$ intervallumon nem korlátos, tehát a sor nem egyenletesen konvergens. Ha azonban pl. csak a $[0, 2; 1]$ intervallumra szorítkozunk, akkor az $N_\epsilon(x)$ függvénynek van felső korlátja, vagyis ezen az intervallumon a sor egyenletesen konvergens.

Az egyenletes konvergencia fogalma kulcsfontosságú a függvénysorok elméletében. Mielőtt ennek bemutatásához kezdenénk, egy egyszerű kritériumot mutatunk az egyenletes konvergencia eldöntésére.

2.15. Tétel.

Weierstrass tétel

Egyenletes konvergencia kritériuma : Ha az $[a; b]$ intervallumon minden $k \in \mathbf{N}$ esetén $|f_k(x)| \leq u_k$ és a $\sum u_k$ numerikus sor konvergens, akkor a $\sum f_k(x)$ függvénysor az $[a; b]$ intervallumon egyenletesen konvergens.

Bizonyítás:

A $\sum u_k$ numerikus sor konvergenciája miatt tetszőleges $\epsilon > 0$ számhoz található N küszöbindex úgy, hogy $n > N$ és $m > N$ esetén

$$|u_{n+1} + u_{n+2} + \dots + u_m| < \epsilon$$

De akkor a feltételben szereplő egyenlőtlenség és a Cauchy konvergencia kritérium alapján az is igaz, hogy

$$|f_{n+1}(x) + f_{n+2}(x) + \dots + f_m(x)| = |s_m(x) - s_n(x)| < \epsilon$$

minden $x \in [a; b]$ esetén. Tehát a függvénysor abszolút és egyenletesen konvergens.

Most is, mint a numerikus sorok esetén, felvetődik a kérdés, hogy a végtelen összegek milyen feltételek mellett öröklik át a véges összegek tulajdonságait. Most olyan tulajdonságokra gondolunk, mint : folytonos függvények összege folytonos, differenciálható függvények összege differenciálható, integrálható függvények összege integrálható, stb. Azonnal látni fogjuk, hogy lényegében minden az egyenletes konvergencián múlik. Igaz ugyanis a következő tétel:

2.16. Tétel.

Ha az $(x - \delta; x + \delta)$ intervallumon minden $f_k(x)$ függvény folytonos és a $\sum f_k(x)$ függvénysor egyenletesen konvergens, akkor az $s(x) = \sum f_k(x)$ összegfüggvény is folytonos.

Bizonyítás:

Elég azt megmutatni, hogy minden $\epsilon > 0$ számhoz található δ úgy, hogy $|s(x) - s(x')| < \epsilon$, ha $|x - x'| < \delta$, ahol $x, x' \in (x - \delta; x + \delta)$. Ha $s_n(x)$ jelöli a sor n - edik részletösszegét, akkor felírhatjuk a következő egyenlőtlenséget:

$$\begin{aligned} |s(x) - s(x')| &= |s(x) - s_n(x) + s_n(x) - s_n(x') + s_n(x') - s(x')| \leq \\ &\leq |s(x) - s_n(x)| + |s_n(x) - s_n(x')| + |s_n(x') - s(x')| \end{aligned}$$

Bármely $\epsilon > 0$ esetén az egyenletes konvergencia miatt található olyan N szám, hogy ha $n > N$, akkor

$$|s(x) - s_n(x)| < \frac{\epsilon}{3}$$

és

$$|s_n(x') - s(x')| < \frac{\epsilon}{3}$$

továbbá $s_n(x)$ folytonossága miatt létezik olyan δ , hogy $|x - x'| < \delta$ esetén

$$|s_n(x) - s_n(x')| < \frac{\epsilon}{3}$$

Ekkor tehát

$$|s(x) - s(x')| < \epsilon$$

Az alábbi példa illusztrálja az egyenletes folytonosság szükségességét a fenti tételben.

2.10. Példa.

Példa: Tekintsük a $[0; 1]$ intervallumon a

$$1 + (x - 1) + (x^2 - x) + \dots + (x^n - x^{n-1}) + \dots$$

függvénysort, ahol tehát $s_n(x) = x^n$. Az Olvasó könnyen ellenőrizheti, hogy itt a függvénysor minden tagja folytonos, de az összegfüggvény nem, ugyanis

$$s(x) = \begin{cases} 0 & \text{ha } 0 \leq x < 1 \\ 1 & \text{ha } x = 1 \end{cases}$$

2.17. Tétel.

Ha minden $f_k(x)$ függvény integrálható az $[a; b]$ intervallumon és a $\sum f_k(x)$ függvénysor itt *egyenletesen* konvergens, akkor az $s(x) = \sum f_k(x)$ összegfüggvény is integrálható az $[a; b]$ intervallumon és $\int_a^b s(x)dx = \sum_{k=1}^{\infty} \int_a^b f_k(x)dx$

A tétel bizonyítását elhagyjuk. Helyette egy példával illusztráljuk, miért szükséges a feltételben az egyenletes konvergencia.

2.11. Példa.

Tekintsük a $[0; 1]$ intervallumon azt a függvénysort, amelynek n -edik részletösszege

$$s_n(x) = 2nxe^{-nx^2}$$

azaz

$$f_n(x) = s_n(x) - s_{n-1}(x)$$

Könnyen belátható, hogy a függvénysor konvergens, és $s(x) = 0$. Következés képpen

$$\int_0^1 s(x)dx = 0$$

Azonban, ha tagonként integrálunk és n tart végtelenhez, akkor

$$\int_0^1 s_n(x)dx = \int_0^1 2nxe^{-nx^2}dx = \left[-e^{-nx^2}\right]_0^1 = -e^{-n} + 1$$

Ez tart 1-hez, ha $n \rightarrow \infty$.

2.18. Tétel.

Ha az $(x - \delta; x + \delta)$ intervallumon minden $f_k(x)$ függvény differenciálható és a $\sum f_k(x)$ függvény-sor konvergens, és $s(x) = \sum f_k(x)$, valamint a deriváltakból alkotott $\sum f'_k(x)$ függvény-sor egyenletesen konvergens, akkor az $s(x) = \sum f_k(x)$ összegfüggvény is differenciálható, és $s'(x) = \sum f'_k(x)$.

A bizonyítást itt is elhagyjuk. Helyette egy példával illusztráljuk, hogy ha a feltételek közül a deriváltakból alkotott sor egyenletes konvergenciája nem teljesül, akkor a tétel állítása nem igaz.

2.12. Példa.

Tekintsük a $[0; 1]$ intervallumon a következő függvény-sort:

$$x + \left(\frac{x^2}{2} - x\right) + \left(\frac{x^3}{3} - \frac{x^2}{2}\right) + \dots + \left(\frac{x^{n+1}}{n+1} - \frac{x^n}{n}\right) + \dots$$

Az összegfüggvény azonosan zérus. Minden tag deriválható, de a deriváltakból alkotott sor összege nem zérus, ugyanis

$$1 + (x - 1) + (x^2 - x) + \dots + (x^n - x^{n-1}) + \dots = \begin{cases} 0 & \text{ha } 0 \leq x < 1 \\ 1 & \text{ha } x = 1 \end{cases}$$

2.2.1. Hatványsorok

2.13. Definíció.

Legyen x_0 tetszőleges rögzített valós szám. Az

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n \tag{23}$$

függvény-sort *hatványsornak* nevezzük.

2.19. Tétel.

Ha létezik a

$$\lim_{n \rightarrow \infty} \left| \frac{a_n}{a_{n+1}} \right| = R \tag{24}$$

határérték, akkor a $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ hatványsor

$$-R < x < R$$

esetén konvergens.

Bizonyítás:

A sorokra vonatkozó hányadoskritérium alapján a $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ hatványsor konvergens, ha

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}(x - x_0)^{n+1}}{a_n(x - x_0)^n} \right| = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} (x - x_0) \right| = |x - x_0| \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| < 1, \quad (25)$$

azaz, ha

$$|x - x_0| < \lim_{n \rightarrow \infty} \left| \frac{a_n}{a_{n+1}} \right| \quad (26)$$

2.3. Megjegyzés.

A bizonyításból az is látszik, hogy ha

a)

$$\lim_{n \rightarrow \infty} \left| \frac{a_n}{a_{n+1}} \right| = \infty$$

akkor a hatványsor minden valós x esetén konvergens.

b)

Ha $\lim_{n \rightarrow \infty} \left| \frac{a_n}{a_{n+1}} \right| = R$ véges, akkor

$$|x - x_0| > R$$

esetén a hatványsor divergens.

c)

Szokás az R számot a hatványsor *konvergencia-sugarának* nevezni.

d)

Az $|x - x_0| = R$ esetről a tétel nem mond semmit. Ezt az esetet külön kell megvizsgálni.

2.20. Tétel.

Egy hatványsor a konvergencia-tartományán belül abszolút konvergens és egyenletesen konvergens.

Bizonyítás:

.....

Következmény:

Egy hatványsor összegfüggvénye a konvergencia tartomány belsejében folytonos.

Egy hatványsor tagonként deriválható és a deriváltakból álló sor összege az összegfüggvény deriváltja.

2.2.2. Taylor sor

2.14. Definíció.

Ha az $f(x)$ függvény akárhányszor differenciálható az x_0 pontban és környezetében, akkor a

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \quad (27)$$

hatványsort *Taylor-sornak* hívjuk.

Fontos kérdés, hogy a Taylor-sor összegfüggvénye megegyezik-e, illetve milyen feltételek teljesülése esetén egyezik meg az $f(x)$ függvénnyel.

2.15. Definíció.

Az $f(x)$ függvény és Taylor-sorának n -ik részletösszege közötti eltérést *maradéktagnak* nevezük.

$$R_n(x) = f(x) - \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k \quad (28)$$

2.21. Tétel. Lagrange-féle maradéktag.

Ha az $f(x)$ függvény elég sokszor differenciálható az x_0 pontban és környezetében, akkor az $R_n(x)$ maradéktag előállítható a következő formában:

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}, \quad (29)$$

ahol $\xi \in (x_0; x)$.

Bizonyítás:

....

2.22. Tétel.

Ha az akárhányszor differenciálható $f(x)$ függvény deriváltjai egy közös korlát alatt maradnak az $(x_0; x)$ intervallumban, akkor itt

$$R_n(x) \rightarrow 0$$

és ezért a Taylor sor konvergencia és összege $f(x)$.

Bizonyítás:

....

2.16. Definíció.

Az $x_0 = 0$ speciális esetben a Taylor sort szokás MacLaurin sornak nevezni.

2.13. Példa.

Néhány nevezetes függvény MacLaurin sora:

$$\begin{aligned}
 e^x &= \sum_{k=0}^{\infty} \frac{x^k}{k!} \\
 chx &= \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} \\
 shx &= \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!} \\
 \cos x &= \sum_{k=0}^{\infty} (-1)^k \cdot \frac{x^{2k}}{(2k)!} \\
 \sin x &= \sum_{k=0}^{\infty} (-1)^k \cdot \frac{x^{2k+1}}{(2k+1)!} \\
 \frac{1}{1-x} &= \sum_{k=0}^{\infty} x^k \quad \text{ha } |x| < 1 \\
 \frac{1}{1+x} &= \sum_{k=0}^{\infty} (-x)^k \quad \text{ha } |x| < 1 \\
 \ln(1+x) &= \sum_{k=1}^{\infty} (-1)^{k+1} \cdot \frac{x^k}{k} \quad \text{ha } |x| < 1 \\
 \frac{1}{1+x^2} &= \sum_{k=0}^{\infty} (-1)^k \cdot x^{2k} \quad \text{ha } |x| < 1 \\
 arctg(x) &= \sum_{k=0}^{\infty} (-1)^k \cdot \frac{x^{2k+1}}{2k+1} \quad \text{ha } |x| < 1
 \end{aligned} \tag{30}$$

2.2.3. Ortogonális sorok

2.17. Definíció.

Az $[a; b]$ intervallumon értelmezett $f(x)$ és $g(x)$ függvényeket ortogonálisnak nevezük, ha

$$\int_a^b f(x)g(x)dx = 0.$$

2.18. Definíció.

Az $[a; b]$ intervallumon értelmezett $\phi_1(x), \phi_2(x), \dots, \phi_n(x), \dots$ függvényrendszert ortogonálisnak nevezük, ha a benne szereplő függvények páronként ortogonálisak, azaz ha

$$\int_a^b \phi_i(x)\phi_j(x)dx = 0,$$

$i \neq j$ esetén.

2.19. Definíció.

Az $[a; b]$ intervallumon értelmezett $\phi_1(x), \phi_2(x), \dots, \phi_n(x), \dots$ függvényrendszert ortonormáltnak nevezük, ha ortogonális rendszer és minden $\phi_i(x)$ függvény esetén

$$\int_a^b \phi_i^2(x)dx = 1.$$

2.20. Definíció.

Az $[a; b]$ intervallumon értelmezett $f(x)$ függvény ortogonális sorfejtése a $\phi_1(x), \phi_2(x), \dots, \phi_n(x), \dots$ függvényrendszer szerint a

$$\sum_{k=1}^{\infty} c_k \cdot \phi_k(x) \quad (31)$$

függvénysor, ahol a c_k együtthatókat a következő integrál definiálja:

$$c_k = \int_a^b f(x) \phi_k(x) dx. \quad (32)$$

A következő tétel azt mutatja, hogy ha két függvény távolságát a különbségük négyzetének integráljával mérjük, akkor egy $f(x)$ függvény legjobb közelítése a $\phi_1(x), \phi_2(x), \dots, \phi_n(x), \dots$ ortogonális rendszer szerint éppen a sorfejtése:

2.23. Tétel.

Az

$$\int_a^b \left[f(x) - \sum_{k=1}^{\infty} a_k \cdot \phi_k(x) \right]^2 dx \quad (33)$$

integrál akkor a legkisebb, ha

$$a_k = c_k = \int_a^b f(x) \phi_k(x) dx.$$

Bizonyítás:

$$\begin{aligned} & \int_a^b [f(x) - \sum_{k=1}^{\infty} a_k \cdot \phi_k(x)]^2 dx = \\ & = \int_a^b \left[f^2(x) - 2f(x) \sum_{k=1}^{\infty} a_k \phi_k(x) + (\sum_{k=1}^{\infty} a_k \cdot \phi_k(x))^2 \right] dx = \\ & = \int_a^b f^2(x) dx - 2 \sum_{k=1}^{\infty} \int_a^b a_k f(x) \phi_k(x) dx + \int_a^b \sum_{k=1}^{\infty} (a_k)^2 (\phi_k(x))^2 dx + \\ & + \int_a^b \sum_{k=1}^{\infty} 2a_k a_j \phi_k(x) \phi_j(x) dx \end{aligned} \quad (34)$$

Az utolsó tag az ortogonalitás miatt zérus. Továbbá

$$\begin{aligned} & = \int_a^b f^2(x) dx - 2 \sum_{k=1}^{\infty} \int_a^b a_k f(x) \phi_k(x) dx + \int_a^b \sum_{k=1}^{\infty} (a_k)^2 (\phi_k(x))^2 dx = \\ & = \int_a^b f^2(x) dx - 2 \sum_{k=1}^{\infty} a_k \int_a^b f(x) \phi_k(x) dx + \sum_{k=1}^{\infty} (a_k)^2 \int_a^b (\phi_k(x))^2 dx = \\ & = \int_a^b f^2(x) dx - 2 \sum_{k=1}^{\infty} a_k c_k + \sum_{k=1}^{\infty} (a_k)^2 \end{aligned} \quad (35)$$

Ezt a következő képpen átalakítva kapjuk, hogy

$$\begin{aligned} & \int_a^b f^2(x) dx - 2 \sum_{k=1}^{\infty} a_k c_k + \sum_{k=1}^{\infty} (a_k)^2 + \sum_{k=1}^{\infty} (c_k)^2 - \sum_{k=1}^{\infty} (c_k)^2 = \\ & = \int_a^b f^2(x) dx - \sum_{k=1}^{\infty} (c_k)^2 + \sum_{k=1}^{\infty} (a_k - c_k)^2 \end{aligned} \quad (36)$$

Itt az első két tag nem függ az a_k együtthatók megválasztásától. Tehát a kifejezés akkor a legkisebb, ha a harmadik tag zérus, azaz minden $a_k = c_k$.

Bizonyítás nélkül állítjuk a következő tételt:

2.24. Tétel.

Ha a $\sum_{k=1}^{\infty} (c_k)^2$ sor konvergens, akkor a $\sum_{k=1}^{\infty} c_k \phi_k(x)$ sor is konvergens és konvergál az $f(x)$ függvényhez olyan értelemben, hogy az eltérés négyzetintegrálja zérushoz tart.

2.2.4. Fourier sor

Könnyű megmutatni, hogy az

$$\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos x, \frac{1}{\sqrt{\pi}} \sin x, \frac{1}{\sqrt{\pi}} \cos 2x, \frac{1}{\sqrt{\pi}} \sin 2x, \dots, \frac{1}{\sqrt{\pi}} \cos kx, \frac{1}{\sqrt{\pi}} \sin kx, \dots$$

trigonometrikus függvényrendszer ortonormált rendszer egy tetszőleges 2π hosszúságú intervallumon. Ehhez csak a következő integrálokat kell ellenőrizni:

$$\begin{aligned} \int_0^{2\pi} \cos kx dx &= 0 \\ \int_0^{2\pi} \sin kx dx &= 0 \\ \int_0^{2\pi} \sin kx \cos lxdx &= 0 \\ \int_0^{2\pi} \sin kx \sin lxdx &= 0 \quad \text{ha } k \neq l \\ \int_0^{2\pi} \cos kx \cos lxdx &= 0 \quad \text{ha } k \neq l \end{aligned} \quad (37)$$

valamint

$$\begin{aligned} \int_0^{2\pi} \frac{1}{2\pi} dx &= 1 \\ \frac{1}{\pi} \int_0^{2\pi} \cos^2 kx dx &= 1 \\ \frac{1}{\pi} \int_0^{2\pi} \sin^2 kx dx &= 1 \end{aligned} \quad (38)$$

2.21. Definíció.

A fenti trigonometrikus rendszer szerinti sort nevezzük Fourier-sornak. Tehát egy 2π szerint periódikus $f(x)$ függvény Fourier sora:

$$f(x) \approx a_0 + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx), \quad (39)$$

ahol

$$\begin{aligned} a_0 &= \frac{1}{2\pi} \int_0^{2\pi} f(x) dx \\ a_k &= \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx dx \quad \text{ha } k \geq 1 \\ b_k &= \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx dx \end{aligned} \quad (40)$$

Ez az eredmény könnyen általánosítható tetszőleges $2l$ hosszúságú intervallumon periódikus függvény sorfejtésére:

$$f(x) \approx a_0 + \sum_{k=1}^{\infty} \left(a_k \cos \frac{k\pi}{l} x + b_k \sin \frac{k\pi}{l} x \right), \quad (41)$$

ahol

$$\begin{aligned} a_0 &= \frac{1}{2l} \int_0^{2l} f(x) dx \\ a_k &= \frac{1}{l} \int_0^{2l} f(x) \cos \frac{k\pi}{l} x dx \quad \text{ha } k \geq 1 \\ b_k &= \frac{1}{l} \int_0^{2l} f(x) \sin \frac{k\pi}{l} x dx \end{aligned} \quad (42)$$

2.4. Megjegyzés.

Könnyen belátható, hogy páros függvény Fourier sora tiszta koszinuszos sor, páratlan függvény Fourier sora tiszta szinuszos sor.

Két tétel a Fourier sorok konvergenciájáról:

2.25. Tétel.

Ha létezik az $f(x)$ függvény második deriváltja és korlátos, akkor az $f(x)$ függvény Fourier sora konvergens.

2.26. Tétel.

Ha létezik az $f(x)$ függvény jobb és bal oldali határértéke az x helyen (jelölje ezeket $f(x+0)$, illetve $f(x-0)$) és ezek eleget tesznek az alábbi, úgynevezett Lipschitz feltételnek:

$$|f(x+h) - f(x+0)| < c \cdot h$$

$$|f(x-h) - f(x-0)| < c \cdot h$$

ahol c egy alkalmas konstans, akkor az $f(x)$ függvény Fourier sora konvergens és konvergál a jobb és bal oldali határérték számtani közepéhez

$$\frac{f(x+0) + f(x-0)}{2}\text{-höz.}$$

3. Egyenletek, egyenletrendszerek numerikus megoldása

3.1. Bevezetés

Ebben a fejezetben egyenletek és egyenletrendszerek numerikus megoldásával foglalkozunk. Először is egyenletek numerikus megoldására mutatunk néhány módszert.

Legyen $f(x)$ egy valós értékű függvény. Az

$$f(x) = 0 \tag{43}$$

egyenlet megoldását (megoldásait) keressük.

A probléma megoldásának első lépése olyan intervallum megtalálása, amelyben van megoldás, azaz találni szeretnénk olyan a illetve b számot, amelyre $f(a)$ és $f(b)$ ellenkező előjelű. Ekkor az $[a; b]$ intervallumban az (43) egyenletnek van gyöke. Ilyen intervallum megtalálása nem mindig egyszerű és nincs univerzális módszer egy ilyen intervallum előállítására. Lehet grafikus módszerrel keresni alkalmas $[a; b]$ intervallumot, vagy próbálgatással. Speciális esetekben lehet módszeresen keresni.

Ilyen eset pl. az, amikor az $f(x)$ függvény polinom. Erről szól a következő tétel.

3.1. Tétel.

Tekintsük a

$$P(x) = a_0x^n + a_1x^{n-1} + \dots + a_n = 0 \quad (44)$$

egyenletet, ahol $a_0 \neq 0$ és legyen továbbá

$$A = \max\{|a_1|, \dots, |a_n|\}.$$

Ekkor a (44) egyenlet gyökei az origó középpontú

$$R = 1 + \frac{A}{|a_0|}$$

sugarú körön belül vannak.

3.1. Megjegyzés.

A tétel természetesen a komplex gyökökről állítja, hogy a mondott körön belül vannak. Ha vannak valós gyökök, akkor azok a $(-R; +R)$ intervallumban helyezkednek el.

Bizonyítás: Legyen?

Következő probléma annak eldöntése, ha tudjuk, hogy az $[a; b]$ intervallumban van gyök, akkor hány van. Monoton $f(x)$ függvény esetén persze csak egy. Ha $f(x)$ polinom, akkor a gyökök számára a következő becslés adható:

becslés a gyökök számára. előjel szabály?

3.2. Intervallum felezés

Az egyenlet megoldási módszerek egyik hatékony eszköze az intervallum felezési eljárás. Tegyük fel, hogy folytonos $f(x)$ függvény esetén az $f(x) = 0$ egyenletnek pontosan egy gyöke van az $[a_0; b_0]$ intervallumban. Például tudjuk, hogy az $f(x)$ függvény monoton is, valamint $f(a_0)$ és $f(b_0)$ ellenkező előjelű. Ekkor az egyenlet x^* gyökét az alábbi eljárással közelítjük: Számítsuk ki először az $f(\frac{a_0+b_0}{2})$ értéket. Ha ez zérus, akkor készen vagyunk. Ha nem, akkor legyen az $[a_1; b_1]$ az $[a_0; b_0]$ intervallumnak az a fele, amelyikben a gyök van, azaz

$$\begin{aligned} a_1 &= a_0 & \text{és} & & b_1 &= \frac{a_0+b_0}{2}, & \text{ha} & & f(a_0) \cdot f(\frac{a_0+b_0}{2}) < 0 \\ a_1 &= \frac{a_0+b_0}{2} & \text{és} & & b_1 &= b_0, & \text{ha} & & f(\frac{a_0+b_0}{2}) \cdot f(b_0) < 0 \end{aligned}$$

Folytassuk ezt az eljárást. Az n -ik lépésben kapott $[a_n; b_n]$ intervallum tartalmazza a gyököt. Az a_n és b_n értékek alsó illetve felső becslést adnak a gyökre. A gyök közelítésének a hibája kisebb mint

$$\frac{b_n - a_n}{2^n}.$$

3.1. Példa.

ide egy egyszerű példa és ábra

3.3. Iteráció

Most tegyük fel, hogy az $f(x)$ differenciálható függvény olyan, hogy az $f(x) = 0$ egyenletet egy alkalmas ekvivalens átalakítással

$$F(x) = x \quad (45)$$

alakra tudjuk hozni úgy, hogy $|F'(x)| \leq q < 1$ egyenlőtlenség is teljesüljön.

3.2. Tétel.

Tegyük fel, hogy az $F(x)$ függvény értelmezve van az $[a; b]$ intervallumon, annak belsejében differenciálható és $|F'(x)| \leq q < 1$. Továbbá ha $x \in [a; b]$, akkor $F(x) \in [a; b]$ is teljesül. Ekkor az

$$x_{n+1} = F(x_n)$$

iteráció tetszőleges $x_0 \in [a; b]$ értékből indulva konvergens és az $F(x) = x$ egyenlet gyökéhez konvergál.

Bizonyítás:

3.2. Példa.

ide egy egyszerű példa és ábra

3.4. Húrmódszer

Tegyük most fel megint, hogy a folytonos $f(x)$ függvény esetén az $f(x) = 0$ egyenletnek pontosan egy gyöke van az $[a_0; b_0]$ intervallumban. Például tudjuk, hogy az $f(x)$ függvény monoton is, valamint $f(a_0)$ és $f(b_0)$ ellenkező előjelű. Ekkor az egyenlet x^* gyökét az alábbi eljárással közelítjük:

Írjuk fel az $(a; f(a))$ és $(b; f(b))$ pontokon átmenő húr egyenletét. Ahol ez a húr metszi az x -tengelyt -jelölje ezt a pontot x_1 - az lesz a gyök első közelítése. Majd válasszuk az $(a; x_1)$ illetve $(x_1; b)$ intervallumok közül azt, amelyik a gyököt tartalmazza. Ismételjük meg az előző lépéseket addig, amíg a kívánt pontosságot elérjük.

Az eljárás tehát a következő:

Kiindulunk az $a = x_0$, $b = x_1$ értékekből. Az n -ik lépésben felírjuk az $(x_{n-1}; f(x_{n-1}))$ és $(x_n; f(x_n))$ pontokon átmenő húr egyenletét, ami

$$y = \frac{f(x_{n-1}) - f(x_n)}{x_{n-1} - x_n}(x - x_n) + f(x_n)$$

Ennek metszéspontja az x -tengellyel

$$\hat{x} = x_n - f(x_n) \frac{x_{n-1} - x_n}{f(x_{n-1}) - f(x_n)}$$

Most vagy $f(x_{n-1}) \cdot f(\hat{x}) < 0$, vagy $f(\hat{x}) \cdot f(x_n) < 0$. Tartsuk meg azt az intervallumot, amelyikre ez a szorzat negatív. (Ha egyenlő nulla, akkor megtaláltuk a pontos megoldást.) Tehát

$$\begin{array}{ll} x_n = x_{n-1} & \text{és} \quad x_{n+1} = \hat{x}, & \text{ha} \quad f(x_{n-1}) \cdot f(\hat{x}) < 0 \\ & x_{n+1} = \hat{x} & \text{ha} \quad f(\hat{x}) \cdot f(x_n) < 0 \end{array}$$

3.3. Példa.

ide egy egyszerű példa és ábra

3.5. Érintő módszer

Most a gyököt tartalmazó intervallum egyik végpontjában érintőt húzunk az $f(x)$ görbéhez és ennek az x -tengellyel való metszéspontja lesz a következő közelítés. Tegyük fel pl., hogy az $f(x)$ függvény monoton növekvő és konvex. Ekkor mindig a jobb oldali végpontban húzott érintővel közelítünk.

ábra

Az algoritmus tehát:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (46)$$

3.4. Példa.

ide egy egyszerű példa

3.6. Newton módszer

Összekombinálva a húr és érintő módszert alsó, felső közelítést kapunk az egyenlet gyökére.
ábra

3.5. Példa.

ide egy egyszerű példa és ábra

3.7. Lineáris egyenletrendszerek megoldása

Ebben a szakaszban az

$$A\vec{x} = \vec{b}$$

egyenletrendszer megoldásával foglalkozunk, ahol A egy $n \times n$ kvadratikusan mátrix, \vec{x} az ismeretlen, $\vec{b} \neq \vec{0}$ adott vektor.

A megoldási módszerek többsége két nagy csoportba sorolható:

- **Direkt módszerek**

Direkt módszernek nevezünk egy olyan eljárást, amely az adatokat pontosnak feltételezve és a számításokat pontosan elvégezve véges sok lépésben előállítja a megoldást. Többnyire akkor használjuk, ha az együttható mátrix nem túl nagy.

- **Iterációs módszerek**

Az iterációs módszerek esetén egy kezdő értékből kiindulva előállítunk egy sorozatot, amelynek a megoldáshoz kell konvergálnia.

Ezeket a módszereket többnyire akkor használjuk, ha az együttható mátrix nagy méretű és lehetőleg sok zérus elemet tartalmaz. Tipikusan ilyenek lépnek fel parciális differenciálegyenletek közelítő megoldásakor.

A legfontosabb módszereket az alábbi mintapéldán mutatjuk be:

3.6. Példa.

$$\begin{aligned} 6x_1 - 4x_2 + x_3 &= 11 \\ 2x_1 + 5x_2 - 3x_3 &= 13 \\ x_1 + 3x_2 - 8x_3 &= 1 \end{aligned} \tag{47}$$

vagyis

$$\begin{bmatrix} 6 & -4 & 1 \\ 2 & 5 & -3 \\ 1 & 3 & -8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 11 \\ 13 \\ 1 \end{bmatrix} \tag{48}$$

3.7.1. Lineáris egyenletrendszerek direkt megoldása

a) Gauss - Jordan módszer

A módszert a (3.6) mintapéldán keresztül mutatjuk be.

Nyilván elég a mátrixba rendezett együtthatókat leírni, mert ezek tartalmazzák az összes információt.

$$\left[\begin{array}{ccc|c} 6 & -4 & 1 & 11 \\ 2 & 5 & -3 & 13 \\ 1 & 3 & -8 & 1 \end{array} \right]$$

1. lépés. Választunk az első oszlopban egy *generáló elemet*, amelyre egyetlen kikötés, hogy zérustól különböző legyen. Jelen esetben a kényelmesebb számolás kedvéért az első oszlop 3. elemét választjuk.

$$\left[\begin{array}{ccc|c} 6 & -4 & 1 & 11 \\ 2 & 5 & -3 & 13 \\ \boxed{1} & 3 & -8 & 1 \end{array} \right]$$

A bekeretezett generáló elem sorának alkalmas számszorosát hozzáadjuk a többi sorhoz úgy, hogy a generáló elem oszlopában a többi elem zérus legyen. Jelen esetben a 3. sor hatszorosát

kivonjuk az első sorból, valamint a 3. sor kétszeresét kivonjuk a második sorból. Ekkor a következő, az eredetivel ekvivalens egyenletrendszer kapjuk:

$$\left[\begin{array}{ccc|c} 0 & -22 & 49 & 5 \\ 0 & \boxed{-1} & 13 & 11 \\ 1 & 3 & -8 & 1 \end{array} \right]$$

2. lépés. Most megismételjük az első lépést úgy, hogy a 2. oszlopban választunk generáló elemet. Bekereteztük a második oszlop második elemét. A második sort, azaz a második egyenletet, végigosztjuk a generáló elemmel, majd a generáló elem sorának alkalmas számszorosát hozzáadjuk a többi sorhoz úgy, hogy a generáló elem oszlopában a többi elem zérus legyen. Ekkor kapjuk, hogy

$$\left[\begin{array}{ccc|c} 0 & 0 & \boxed{-237} & -237 \\ 0 & 1 & -13 & -11 \\ 1 & 0 & 31 & 34 \end{array} \right]$$

3. lépés. Megismételjük az eljárást a 3. oszlopban bekeretezett elemmel. Ekkor kapjuk, hogy

$$\left[\begin{array}{ccc|c} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 2 \\ 1 & 0 & 0 & 3 \end{array} \right]$$

Ezzel az eljárás befejeződött. A mátrixból kiolvastva az egyenleteket kapjuk a megoldást:

$$x_1 = 3, \quad x_2 = 2, \quad x_3 = 1.$$

b) Trianguláris felbontás módszere

A módszer lényege, hogy az \mathbf{A} együttható mátrixot felbontjuk egy alsó és egy felső háromszög-mátrix szorzatára:

$$\mathbf{A} = \mathbf{LU},$$

ahol

$$\mathbf{L} = \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ l_{31} & l_{32} & 1 & \\ \vdots & & & \\ l_{n1} & l_{n2} & \cdots & 1 \end{bmatrix}$$

és a főátló felett ki nem töltött helyeken minden elem zérus. Továbbá

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \\ & & & u_{nn} \end{bmatrix}$$

és a főátló alatt ki nem töltött helyeken minden elem zérus.

Az egyenletrendszer megoldása majd két lépésben történik. Először megoldjuk az

$$\mathbf{L}\vec{y} = \vec{b}$$

egyenletrendszert, majd az

$$\mathbf{U}\vec{x} = \vec{y}$$

egyenletrendszert. Az így kapott megoldás természetesen megoldása az eredeti problémának, mert

$$\mathbf{A}\vec{x} = \mathbf{L}\mathbf{U}\vec{x} = \mathbf{L}\vec{y} = \vec{b}.$$

Kimutatható, hogy ezen módszer műveletigénye a Gauss-módszer műveletigényével azonos nagyságrendű; mindkettő n^3 .

A trianguláris felbontás a *minimális diadikus felbontással* történhet.

3.1. Definíció.

Diádnak nevezzük egy oszlopvektor és egy sorvektor (ilyen sorrendben vett) szorzatát.

3.7. Példa.

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} [4, 5, 6] = \begin{bmatrix} 4 & 5 & 6 \\ 8 & 10 & 12 \\ 12 & 15 & 18 \end{bmatrix}$$

3.8. Példa.

A minimális diadikus felbontás algoritmus:

Minden lépésben leválasztunk egy diádot a redukálással előállított mátrixból. Az eljárást most is a (3.6) mintapéldán keresztül mutatjuk be.

1. lépés.

$$\begin{aligned} \mathbf{A}^{(1)} &= \mathbf{A} - \frac{1}{6} \begin{bmatrix} 6 \\ 2 \\ 1 \end{bmatrix} [6, -4, 1] = \mathbf{A} - \begin{bmatrix} 6 & -4 & 1 \\ 2 & -\frac{4}{3} & \frac{1}{3} \\ 1 & -\frac{2}{3} & \frac{1}{6} \end{bmatrix} = \\ &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{19}{3} & -\frac{10}{3} \\ 0 & \frac{11}{3} & -\frac{49}{3} \end{bmatrix} \end{aligned}$$

2. lépés.

Megismételjük a diád leválasztást.

$$\begin{aligned} \mathbf{A}^{(2)} &= \mathbf{A}^{(1)} - \frac{3}{19} \begin{bmatrix} 0 \\ \frac{19}{3} \\ \frac{11}{3} \end{bmatrix} [0, \frac{19}{3}, -\frac{10}{3}] = \mathbf{A}^{(1)} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{19}{3} & -\frac{10}{3} \\ 0 & \frac{11}{3} & -\frac{110}{57} \end{bmatrix} = \\ &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\frac{237}{38} \end{bmatrix} \end{aligned}$$

Következő képpen az \mathbf{A} mátrix felbontása:

$$\mathbf{A} = \begin{bmatrix} 6 & -4 & 1 \\ 2 & 5 & -3 \\ 1 & 3 & -8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{1}{6} & \frac{11}{19} & 1 \end{bmatrix} \begin{bmatrix} 6 & -4 & 1 \\ 0 & \frac{19}{3} & -\frac{10}{3} \\ 0 & 0 & -\frac{237}{38} \end{bmatrix}$$

Ennek alapján a (3.6) egyenletrendszer megoldása úgy történik, hogy először megoldjuk az

$$\mathbf{L}\vec{y} = \vec{b}$$

azaz

$$\begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ \frac{1}{6} & \frac{11}{19} & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 11 \\ 13 \\ 1 \end{bmatrix}$$

egyenletrendszert. Itt könnyű felülről lefelé meghatározni az ismeretleneket:

$$\begin{aligned} y_1 &= 11 \\ y_2 &= 13 - \frac{1}{3}y_1 = 13 - \frac{11}{3} = \frac{28}{3} \\ y_3 &= 1 - \frac{1}{6}y_1 - \frac{11}{19}y_2 = 1 - \frac{11}{6} - \frac{11}{19} \cdot \frac{28}{3} = -\frac{237}{38} \end{aligned}$$

Következik a

$$\begin{bmatrix} 6 & -4 & 1 \\ 0 & \frac{19}{3} & -\frac{10}{3} \\ 0 & 0 & -\frac{237}{38} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 11 \\ \frac{28}{3} \\ -\frac{237}{38} \end{bmatrix}$$

egyenletrendszer megoldása. Itt a harmadik egyenletből azonnal látszik, hogy $x_3 = 1$. A második egyenletből

$$x_2 = \frac{3}{19} \left(\frac{28}{3} + \frac{10}{3} \right) = 2.$$

Végül

$$x_1 = \frac{1}{6} (11 + 4 \cdot 2 - 1) = 3.$$

3.7.2. Lineáris egyenletrendszerek iterációs megoldása

A feladat most is az

$$\mathbf{A}\vec{x} = \vec{b}$$

egyenletrendszer megoldása. Tegyük fel, hogy ez az egyenletrendszer ekvivalens átalakítással az

$$\vec{x} = \mathbf{B}\vec{x} + \vec{r} \quad (49)$$

alakra hozható ahol a \mathbf{B} mátrixra majd bizonyos feltételt írunk elő. Bontsuk fel továbbá a \mathbf{B} mátrixot két részre:

$$\mathbf{B} = \mathbf{B}_1 + \mathbf{B}_2. \quad (50)$$

Használni fogjuk a

$$\vec{x}^{(k+1)} = \mathbf{B}_1\vec{x}^{(k+1)} + \mathbf{B}_2\vec{x}^{(k)} + \vec{r} \quad (51)$$

iterációt adott $\vec{x}^{(0)}$ kezdőérték esetén.

Az könnyen látható, hogy ha az $\vec{x}^{(k)}$ sorozat konvergens és $\vec{x}^{(k)} \rightarrow \vec{x}^*$, akkor \vec{x}^* kielégíti a (49) egyenletet és így megoldása az eredeti problémának.

Az alkalmazhatóság szempontjából két alapvető kérdést kell megvizsgálni:

- Milyen feltételek mellett lesz az iterációs eljárás bármely $\vec{x}^{(0)}$ kezdővektor esetén konvergens?
- Hogyan lehet megállapítani azt, hogy már kellő pontossággal közelítettük az igazi megoldást, vagyis hogyan lehet becsülni az $\vec{x}^{(k)}$ és \vec{x}^* vektorok eltérését?

Intermezzo

A folytatáshoz szükségünk lesz a vektor- és mátrixnorma fogalmára.

3.2. Definíció.

Egy \vec{x} vektor normája - jelölje ezt $\|\vec{x}\|$ - egy a vektorhoz rendelt olyan szám, amelyre a következő feltételek teljesülnek:

- 1) $\|\vec{x}\| \geq 0$ és $\|\vec{x}\| = 0$ akkor és csak akkor teljesül, ha $\vec{x} = \vec{0}$.
- 2) $\|\lambda\vec{x}\| = |\lambda| \cdot \|\vec{x}\|$ minden komplex λ szám esetén.
- 3) $\|\vec{x} + \vec{y}\| \geq \|\vec{x}\| + \|\vec{y}\|$ háromszög-egyenlőtlenség.

Leggyakrabban a következő normákat használják:

- a) $\|\vec{x}\|_1 = \sum_{i=1}^n |x_i|$
- b) $\|\vec{x}\|_2 = [\sum_{i=1}^n |x_i|^2]^{1/2}$ Euklideszi norma
- c) $\|\vec{x}\|_\infty = \max_i |x_i|$

3.2. Megjegyzés.

Nem nehéz ellenőrizni, hogy mindhárom norma kielégíti a (3.2) definícióban megadott feltételeket.

Egy mátrix normájától a következő tulajdonságok teljesülését várjuk el:

- 1) $\|\mathbf{A}\| \geq 0$ és $\|\mathbf{A}\| = 0$ akkor és csak akkor teljesül, ha $\mathbf{A} = \mathbf{0}$.
- 2) $\|\lambda\mathbf{A}\| = |\lambda| \cdot \|\mathbf{A}\|$ minden komplex λ szám esetén.
- 3) $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$ háromszög-egyenlőtlenség.
- 4) $\|\mathbf{A} \cdot \mathbf{B}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{B}\|$ Cauchy-Schwarz egyenlőtlenség.

A mátrixnormát a vektornormák segítségével definiáljuk. Ezek a vektornorma által indukált normák.

3.3. Definíció.

Egy \mathbf{A} mátrix normája a

$$\|\mathbf{A}\| = \sup_{\|\vec{x}\| \neq 0} \frac{\|\mathbf{A}\vec{x}\|}{\|\vec{x}\|} \quad (52)$$

3.3. Megjegyzés.

Itt is könnyű ellenőrizni, hogy az így definiált mátrixnorma kielégíti az elvárt feltételeket.

3.4. Megjegyzés.

A mátrixnorma 2) tulajdonsága alapján az (3.3) definíció az alábbi ekvivalens alakba is írható:
Egy \mathbf{A} mátrix normája a

$$\|\mathbf{A}\| = \sup_{\|\vec{x}\|=1} \|\mathbf{A}\vec{x}\| \quad (53)$$

3.3. Tétel.

A fent említett három vektornorma által indukált mátrixnormák az alábbi módon kiszámíthatók:

-

$$\|\mathbf{A}\|_1 = \max_{j=1, \dots, n} \sum_{i=1}^n |a_{ij}|$$

-

$$\|\mathbf{A}\|_2 = \max_{j=1, \dots, n} [\lambda_j \mathbf{A}^* \mathbf{A}]^{1/2},$$

ahol λ_j az $\mathbf{A}^* \mathbf{A}$ mátrix sajátértékein fut végig.

-

$$\|\mathbf{A}\|_\infty = \max_{i=1, \dots, n} \sum_{j=1}^n |a_{ij}|$$

Intermezzo vége

A lineáris egyenletrendszerek iterációs megoldásának folytatása

3.4. Tétel.

Az

$$\vec{x}^{(k+1)} = \mathbf{B}\vec{x}^{(k)} + \vec{r} \quad (54)$$

iteráció akkor és csak akkor konvergens tetszőleges $\vec{x}^{(0)}$ kezdővektor esetén, ha a \mathbf{B} mátrix minden sajátértéke az egységkör belsejében van.

3.5. Tétel.

Az

$$\vec{x}^{(k+1)} = \mathbf{B}_1 \vec{x}^{(k+1)} + \mathbf{B}_2 \vec{x}^{(k)} + \vec{r} \quad (55)$$

iteráció $\|\mathbf{B}\| < 1$ esetén konvergál az egyenletrendszer \vec{x}^* megoldásához és

$$\|\vec{x}^{(k+1)} - \vec{x}^*\| \leq \frac{\|\mathbf{B}_2\|}{1 - \|\mathbf{B}\|} \cdot \|\vec{x}^{(k+1)} - \vec{x}^{(k)}\|.$$

3.5. Megjegyzés.

Ez utóbbi eredmény érvényes az

$$\vec{x}^{(k+1)} = \mathbf{B} \vec{x}^{(k)} + \vec{r}$$

iterációra is $\mathbf{B}_1 = \mathbf{0}$ és $\mathbf{B}_2 = \mathbf{B}$ választással. Ekkor

$$\|\vec{x}^{(k+1)} - \vec{x}^*\| \leq \frac{\|\mathbf{B}\|}{1 - \|\mathbf{B}\|} \cdot \|\vec{x}^{(k+1)} - \vec{x}^{(k)}\|.$$

3.7.3. Relaxációs módszerek lineáris egyenletrendszerek iterációs megoldására

Továbbra is az

$$\mathbf{A} \vec{x} = \vec{b}$$

egyenletrendszer iterációs megoldásával foglalkozunk. A könnyebb áttekinthetőség kedvéért a módszert egy kisméretű példán mutatjuk be. Legyen a megoldandó egyenletrendszer:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned} \quad (56)$$

vagyis

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad (57)$$

Ha az \mathbf{A} mátrix főátlóbeli elemei zérustól különböznek, akkor kifejezhetjük mindegyik egyenletből a megfelelő ismeretlent. Így adódik, hogy

$$\begin{aligned} x_1 &= e_{12}x_2 + e_{13}x_3 + c_1 \\ x_2 &= e_{21}x_1 + e_{23}x_3 + c_2 \\ x_3 &= e_{31}x_1 + e_{32}x_2 + c_3 \end{aligned} \quad (58)$$

ahol

$$e_{ij} = \begin{cases} -a_{ij}/a_{ii} & \text{ha } i \neq j \\ 0 & \text{ha } i = j \end{cases} \quad (59)$$

és

$$c_i = b_i/a_{ii}.$$

Mátrixos alakban:

$$\vec{x} = \mathbf{E}\vec{x} + \vec{c},$$

ahol

$$\mathbf{E} = \mathbf{D}^{-1}\mathbf{C}$$

$$\vec{c} = \mathbf{D}^{-1}\vec{b}$$

és

$$\mathbf{D} = \text{diag}\mathbf{A}, \quad \mathbf{C} = \mathbf{D} - \mathbf{A}.$$

I. Jacobi-módszer

Tetszőleges $\vec{x}^{(0)} = [x_1^{(0)}; x_2^{(0)}; x_3^{(0)}]^T$ értékekből indulva kiszámítjuk az $\vec{x}^{(1)}$ vektor koordinátáit úgy, hogy $\vec{x}^{(0)}$ koordinátáit behelyettesítjük a (58) egyenletrendszer jobb oldalába. A kapott $\vec{x}^{(1)}$ vektorból hasonlóan $\vec{x}^{(2)}$ vektort, és így tovább. Általánosan:

$$\begin{aligned} x_1^{(n+1)} &= e_{12}x_2^{(n)} + e_{13}x_3^{(n)} + c_1 \\ x_2^{(n+1)} &= e_{21}x_1^{(n)} + e_{23}x_3^{(n)} + c_2 \\ x_3^{(n+1)} &= e_{31}x_1^{(n)} + e_{32}x_2^{(n)} + c_3 \end{aligned} \quad (60)$$

Ugyan ez mátrix alakban:

$$\vec{x}^{(n+1)} = \mathbf{E}\vec{x}^{(n)} + \vec{c}.$$

A Jacobi-módszert is a (3.6) mintapéldán mutatjuk be:

3.9. Példa.

Megoldandó az alábbi egyenletrendszer Jacobi-módszerrel.

$$\begin{aligned} 6x_1 - 4x_2 + x_3 &= 11 \\ 2x_1 + 5x_2 - 3x_3 &= 13 \\ x_1 + 3x_2 - 8x_3 &= 1 \end{aligned} \quad (61)$$

Először átrendezzük az egyenletrendszert az (58) alaknak megfelelően:

$$\begin{aligned} x_1 &= \frac{4}{6}x_2 - \frac{1}{6}x_3 + \frac{11}{6} \\ x_2 &= -\frac{2}{5}x_1 + \frac{3}{5}x_3 + \frac{13}{5} \\ x_3 &= \frac{1}{8}x_1 + \frac{3}{8}x_2 - \frac{1}{8} \end{aligned} \quad (62)$$

Az

$$\begin{aligned} x_1^{(n+1)} &= \frac{4}{6}x_2^{(n)} - \frac{1}{6}x_3^{(n)} + \frac{11}{6} \\ x_2^{(n+1)} &= -\frac{2}{5}x_1^{(n)} + \frac{3}{5}x_3^{(n)} + \frac{13}{5} \\ x_3^{(n+1)} &= \frac{1}{8}x_1^{(n)} + \frac{3}{8}x_2^{(n)} - \frac{1}{8} \end{aligned} \quad (63)$$

iterációt indítsuk az $x_1^{(0)} = 0$, $x_2^{(0)} = 0$, $x_3^{(0)} = 0$ pontból. Ekkor a következő sorozatot kapjuk:

1. táblázat. Jacobi-módszer

n	1	2	3	4	5	6	7	8	9	10
$x_1^{(n)}$	1,8333	3,5875	2,8479	2,8757	3,0535	2,9963	2,9873	3,0042	3,0005	2,9987
$x_2^{(n)}$	2,6000	1,7916	1,8125	2,0580	1,9960	1,9823	2,0045	2,0008	1,9983	2,0002
$x_3^{(n)}$	-0,1250	1,0791	0,9953	0,9106	1,0062	1,0052	0,9929	1,0001	1,0008	0,9994

Látjuk, hogy az iteráció gyorsan konvergál a megoldáshoz.

II. Relaxációs módszer

Az előző eljárást úgy módosítjuk, hogy választunk egy $w \neq 0$ valós paramétert és a következő iterációt hajtjuk végre:

$$\begin{aligned} x_1^{(n+1)} &= w(e_{12}x_2^{(n)} + e_{13}x_3^{(n)} + c_1) + (1-w)x_1^{(n)} \\ x_2^{(n+1)} &= w(e_{21}x_1^{(n)} + e_{23}x_3^{(n)} + c_2) + (1-w)x_2^{(n)} \\ x_3^{(n+1)} &= w(e_{31}x_1^{(n)} + e_{32}x_2^{(n)} + c_3) + (1-w)x_3^{(n)} \end{aligned} \quad (64)$$

Ez ekvivalens mátrix alakban:

$$\vec{x}^{(n+1)} = \mathbf{E}_w \vec{x}^{(n)} + w\vec{c}.$$

ahol

$$\mathbf{E}_w = w\mathbf{E} + (1-w)\mathbf{I}.$$

Vegyük észre, hogy ha $w = 1$, akkor visszakapjuk a Jacobi-módszert.

Ha $w > 1$, akkor azt mondjuk, hogy a módszer *túlrelaxált*.

Ha $w < 1$, akkor módszer *alulrelaxált*.

A w paraméter megválasztásától függ a konvergencia sebessége.

III. Gauss-Seidel módszer

Ez annyiban tér el a Jacobi-módszertől, hogy az $\vec{x}^{(n+1)}$ vektor koordinátáit már hamarabb felhasználjuk. A (60) iteráció így módosul:

$$\begin{aligned} x_1^{(n+1)} &= e_{12}x_2^{(n)} + e_{13}x_3^{(n)} + c_1 \\ x_2^{(n+1)} &= e_{21}x_1^{(n+1)} + e_{23}x_3^{(n)} + c_2 \\ x_3^{(n+1)} &= e_{31}x_1^{(n+1)} + e_{32}x_2^{(n+1)} + c_3 \end{aligned} \quad (65)$$

Mátrixos alakban:

$$\vec{x}^{(n+1)} = \mathbf{L}\vec{x}^{(n+1)} + \mathbf{U}\vec{x}^{(n)} + \vec{c},$$

ahol \mathbf{L} alsó, \mathbf{U} felső háromszög-mátrix:

$$\mathbf{L} = \begin{bmatrix} 0 & 0 & 0 \\ e_{21} & 0 & 0 \\ e_{31} & e_{32} & 0 \end{bmatrix} \quad \mathbf{U} = \begin{bmatrix} 0 & e_{12} & e_{13} \\ 0 & 0 & e_{23} \\ 0 & 0 & 0 \end{bmatrix} \quad (66)$$

3.10. Példa.

A Gauss-Seidel-módszer alkalmazása a (3.6) mintapéldára:

n	1	2	3	4	5	6	7	8	9	10
$x_1^{(n)}$	1,8333	2,9437	2,9437	2,9992	2,9972	3,0000	3,0000	2,9998	3,0000	2,9999
$x_2^{(n)}$	1,8666	1,9050	1,9968	1,9953	2,0000	1,9997	1,9997	2,0000	1,9999	2,0000
$x_3^{(n)}$	0,8041	0,9573	0,9918	0,9981	0,9996	0,9999	0,9999	0,9999	0,9999	0,9999

2. táblázat. Gauss-Seidel-módszer

Az iteráció gyorsabban konvergál, mint a Jacobi-módszer esetén.

IV. Szukcesszív túlrelaxálás módszere

Ez annyiban tér el a relaxációs-módszertől, hogy az $\vec{x}^{(n+1)}$ vektor koordinátáit már hamarabb felhasználjuk. Az iteráció így módosul:

$$\begin{aligned}
 x_1^{(n+1)} &= w(e_{11}x_1^{(n+1)} + e_{12}x_2^{(n)} + e_{13}x_3^{(n)} + c_1) + (1-w)x_1^{(n)} \\
 x_2^{(n+1)} &= w(e_{21}x_1^{(n+1)} + e_{22}x_2^{(n+1)} + e_{23}x_3^{(n)} + c_2) + (1-w)x_2^{(n)} \\
 x_3^{(n+1)} &= w(e_{31}x_1^{(n+1)} + e_{32}x_2^{(n+1)} + e_{33}x_3^{(n+1)} + c_3) + (1-w)x_3^{(n)}
 \end{aligned} \tag{67}$$

Ez ekvivalens mátrix alakban:

$$\vec{x}^{(n+1)} = w(\mathbf{L}\vec{x}^{(n+1)} + \mathbf{U}\vec{x}^{(n)} + \vec{c}) + (1-w)\vec{x}^{(n)}.$$

4. Közösleges elsőrendű differenciálegyenletek numerikus megoldása

Műszaki, fizikai, mechanikai problémák matematikai modellezésének természetes eszköze a jelenség differenciálegyenletekkel való leírása. A differenciálegyenlet olyan egyenlet, amelyben az ismeretlen egy függvény és az egyenlet tartalmazza az ismeretlen függvény deriváltját, vagy deriváltjait.

Ha az ismeretlen függvény egyváltozós, akkor a differenciálegyenletet *közösleges differenciálegyenletnek* nevezzük. Ha az ismeretlen függvény többváltozós és az egyenletben annak parciális deriváltjai szerepelnek, akkor a differenciálegyenletet *parciális differenciálegyenletnek* hívjuk. Néhány egyszerű példán mutatjuk be, hogyan vezet egy mechanikai probléma differenciálegyenletre.

4.1. Példa.

Tegyük fel, hogy egy "láncíd" típusú híd esetén a tartókötél súlya elhanyagolható az útpálya súlyához képest. A felfüggesztés sűrű, az útpálya egyenletes tömegeloszlású. Milyen lesz a tartókötél alakja?

kell ábra !!!

megoldás:

Függőleges és vízszintes erők egyensúlya a szimmetria tengelytől x távolságra:

$$F \sin \phi = kx \quad F \cos \phi = H$$

Ebből

$$\operatorname{tg} \phi = y'(x) = \frac{kx}{H} \quad (68)$$

Mint látjuk ez a probléma egy differenciálegyenletre vezet. Mivel ez az egyenlet az ismeretlen $y(x)$ függvény első deriváltját tartalmazza, ezért ezt *közönséges elsőrendű differenciálegyenletnek* nevezzük.

Ezt a differenciálegyenletet könnyen meg tudjuk oldani integrálással és kapjuk, hogy

$$y(x) = \frac{k}{2H}x^2 + c$$

Tehát a kötélek alakja parabola. A c integrációs konstans értékét az szabja meg, hogyan helyeztük el a modellt a koordinátarendszerben.

4.2. Példa.

Tegyük most fel, hogy egy "láncid" típusú híd esetén a tartókötel súlya jelentős, de az útpálya súlya elhanyagolható. A kötel egyenletes tömegeloszlású. Milyen lesz a tartókötel alakja? *ide is kell ábra !!!*

Megoldás:

Függőleges és vízszintes erők egyensúlya a szimmetria tengelytől x távolságra:

$$F \sin \phi = ks(x) \quad F \cos \phi = H,$$

ahol $s(x)$ a kötél darab hossza. Innen

$$\operatorname{tg} \phi = y'(x) = \frac{k}{H}s(x) \quad (69)$$

Mivel az ívhossz deriváltját tudjuk kifejezni

$$\frac{ds}{dx} = \sqrt{1 + (y')^2},$$

ezért deriváljuk a (69) egyenletet.

$$y''(x) = \frac{k}{H} \sqrt{1 + (y')^2} \quad (70)$$

Most tehát egy *közönséges másodrendű differenciálegyenletet* kaptunk. Ezt azon feltételek mellett kell megoldani, hogy $y(0) = \text{adott}$ és $y'(0) = 0$.

Az egyenlet explicite megoldható. A megoldás:

$$y(x) = \frac{H}{k} \operatorname{ch} \frac{k}{H} x + c$$

vagyis a súlyos kötel alakja koszinusz-hiperbolikus függvénnyel írható le.

4.1. Megjegyzés.

Ha egy híd esetén a kötélsúlyát is, és az útpálya súlyát is figyelembe vesszük, akkor a kapott másodrendű differenciálegyenletet nem tudjuk explicite megoldani. Ilyenkor csak a numerikus módszerek segítenek.

4.1. Szukcesszív approximáció

Először olyan közönséges elsőrendű differenciálegyenletekkel foglalkozunk, amelyeknek általános alakja

$$y'(x) = f(x, y(x)).$$

Nagyon fontos kérdés, hogy van-e egyáltalán megoldása egy ilyen egyenletnek, illetve milyen feltételek mellett van? Műszaki szempontból továbbá az is alapvetően fontos, hogy tudjuk hány megoldás van. Erre vonatkozik az alábbi definíció és tétel.

4.1. Definíció.

Az $y'(x) = f(x, y(x))$ differenciálegyenletet a hozzá tartozó $y(x_0) = y_0$ kezdeti feltétellel *kezdetiérték-problémának* vagy *Cauchy-feladatnak* hívják.

4.1. Tétel.

Az

$$y'(x) = f(x, y(x)) \quad y(x_0) = y_0 \tag{71}$$

kezdetiérték-problémának van megoldása ha, az $f(x, y)$ kétváltozós függvény folytonos az $(x_0; y_0)$ pontban és annak egy környezetében. Továbbá egyetlen megoldása van, ha az $f(x, y)$ függvény kielégíti az $(x_0; y_0)$ pont környezetében az $|f(x, y_1) - f(x, y_2)| < k|y_1 - y_2|$ úgynevezett Lipschitz-feltételt is.

A bizonyítás - amit itt nem írunk le részletesen - a *szukcesszív approximációra* épül. Ennek gondolatmenete a következő:

Először is írjuk át a kezdeti érték problémát egy integrálegyenlet alakra:

$$y(x) = y_0 + \int_{x_0}^x f(t, y(t)) dt \tag{72}$$

Deriválással könnyen ellenőrizhető, hogy ez ekvivalens a (71) kezdetiérték- problémával. A sukceszív approximáció valójában a (72) egyenlet iterációs megoldása:

$$y_{n+1}(x) = y_0 + \int_{x_0}^x f(t, y_n(t)) dt \tag{73}$$

Az iteráció konvergenciáját a (4.1) tételben megfogalmazott feltételek biztosítják.

4.3. Példa.

Közelítsük az $y' = x - y$, $y(0) = 1$ kezdetiérték-probléma megoldását szukcesszív approximációval.

Megoldás: Először is látjuk, hogy az $f(x, y) = x - y$ függvény folytonos, másrészt $|(x - y_1) - (x - y_2)| = |y_1 - y_2|$ ezért kielégíti a Lipschitz-feltételt is $k = 1$ konstans választással.

Kezdjük a közelítést az $y_0 = 1$ függvénnyel. Ekkor

$$y_1 = 1 + \int_0^x (t - 1)dt = 1 + \left[\frac{1}{2}t^2 - t\right]_0^x = \frac{1}{2}x^2 - x + 1$$

$$y_2 = 1 + \int_0^x \left(t - \left(\frac{1}{2}t^2 - t + 1\right)\right)dt = -\frac{1}{6}x^3 + x^2 - x + 1$$

$$y_3 = 1 + \int_0^x \left(t - \left(-\frac{1}{6}t^3 + t^2 - t + 1\right)\right)dt = \frac{1}{24}x^4 - \frac{1}{3}x^3 + x^2 - x + 1$$

$$y_4 = -\frac{1}{120}x^5 + \frac{1}{12}x^4 - \frac{1}{3}x^3 + x^2 - x + 1$$

$$y_5 = \frac{1}{720}x^6 - \frac{1}{60}x^5 + \frac{1}{12}x^4 - \frac{1}{3}x^3 + x^2 - x + 1$$

⋮

4.2. Megjegyzés.

A (4.3) példabeli kezdetiérték-probléma megoldható expliciten is, és a megoldása $y = x - 1 + 2e^{-x}$. A szukcesszív approximációval kapott közelítés a megoldásfüggvény megfelelő Taylor-polinomjától mindig csak egyetlen tagban tér el.

4.2. Euler módszer

Ez az egyik legegyszerűbb módszer a

$$y'(x) = f(x, y(x)), \quad y(x_0) = y_0$$

kezdetiérték-probléma megoldásának közelítésére. Keressük a megoldást egy olyan $[a; b]$ intervallumon, ahol $a = x_0$. Osszuk fel az $[a; b]$ intervallumot m részre. Legyenek az osztópontok

$$a = x_0 < x_1 < \dots < x_m = b$$

A megoldásfüggvényt az osztópontokban fogjuk közelíteni, az osztópontok között pedig egyes szakasszal kötjük össze a közelítő pontokat.

ábra!!!

Jelölje a közelítő függvényt \hat{y} . Természetesen $\hat{y}(x_0) = y_0$. Induljon a töröttvonal érintő irányában. Ennek meredeksége éppen a differenciálhányados, ezért az $(x_0; x_1)$ intervallumon a közelítés: $y = f(x_0, y_0)(x - x_0) + y_0$. Tehát az x_1 pontban

$$\hat{y}(x_1) = f(x_0, y_0)(x_1 - x_0) + y_0$$

Most az x_1 pontban $f(x_1, \hat{y}(x_1))$ fogja közelíteni a differenciálhányadost, ezért

$$\hat{y}(x_2) = f(x_1, \hat{y}(x_1))(x_2 - x_1) + \hat{y}(x_1).$$

És így tovább

$$\hat{y}(x_{k+1}) = f(x_k, \hat{y}(x_k))(x_{k+1} - x_k) + \hat{y}(x_k).$$

4.4. Példa.

Oldjuk meg újra az előző (4.3) példát

$$y' = x - y, \quad y(0) = 1$$

most Euler módszerrel. Tegyük fel, hogy a $[0; 0,25; 0,5; 0,75; 1]$ pontokban van szükségünk a megoldás közelítő értékére.

Megoldás: A $[0; 0,25]$ intervallumon az

$$y = f(x_0, y_0)(x_1 - x_0) + y_0 = -x + 1$$

közelítést használjuk, ezért $x_1 = 0,25$ -nél a közelítés $\hat{y}_1 = 0,75$.

A második intervallumon:

$$\hat{y}(x_2) = f(x_1, \hat{y}_1)(x_2 - x_1) + \hat{y}_1 = -0,5 \cdot 0,25 + 0,75 = 0,625.$$

A harmadik intervallumon:

$$\hat{y}(x_3) = f(x_2, \hat{y}_2)(x_3 - x_2) + \hat{y}_2 = -0,125 \cdot 0,25 + 0,625 = 0,59375.$$

Végül a negyedik intervallumon:

$$\hat{y}(x_4) = f(x_3, \hat{y}_3)(x_4 - x_3) + \hat{y}_3 = 0,15625 \cdot 0,25 + 0,59375 = 0,6328125.$$

Összevetve a

$$0,75; \quad 0,625 \quad 0,59375 \quad 0,6328125$$

közelítő értékeket a pontos $y = x - 1 + 2e^{-x}$ megoldásból származó

$$0,8076 \quad 0,7130 \quad 0,6947 \quad 0,7357$$

értékekkel látjuk, hogy az eltérés legfeljebb 0,1.

4.3. Runge - Kutta módszer

Az Euler-módszerhez hasonlóan most is az

$$y'(x) = f(x, y(x)), \quad y(x_0) = y_0$$

kezdetiérték-probléma megoldását közelítjük az

$$a = x_0 < x_1 < \dots < x_m = b$$

osztópontokban. Az osztópontok között egyenes szakasszal kötjük össze a közelítő pontokat. Jelölje h_i az osztópontok közötti távolságot, azaz

$$h_i = x_{i+1} - x_i.$$

Runge - Kutta módszer szerint a közelítés:

$$\begin{aligned} \hat{y}(x_0) &= y_0 \\ &\vdots \\ \hat{y}(x_{i+1}) &= \hat{y}(x_i) + \frac{h_i}{6}(k_1 + 2k_2 + 2k_3 + k_4) \end{aligned}$$

ahol

$$\begin{aligned} k_1 &= f(x_i, \hat{y}(x_i)) \\ k_2 &= f\left(x_i + \frac{h_i}{2}, \hat{y}(x_i) + \frac{h_i}{2}k_1\right) \\ k_3 &= f\left(x_i + \frac{h_i}{2}, \hat{y}(x_i) + \frac{h_i}{2}k_2\right) \\ k_4 &= f(x_i + h_i, \hat{y}(x_i) + h_ik_3) \end{aligned}$$

4.5. Példa.

Oldjuk meg újra az (4.3) példát

$$y' = x - y, \quad y(0) = 1$$

most Runge-Kutta módszerrel. Tegyük fel, hogy a $[0; 0,25; 0,5; 0,75; 1]$ pontokban van szükségünk a megoldás közelítő értékére.

Megoldás:

i=	0	1	2	3	4
x_i	0	0,25	0,5	0,75	1
k_1		-0,75	-0,362915039	-0,061449945	0,173333661
k_2		-0,53125	-0,192550659	0,071231298	0,276666953
k_3		-0,55859375	-0,213846207	0,054646143	0,263750292
k_4		-0,360351563	-0,059453487	0,174888519	0,357396088
\hat{y}_i	1	0,862915039	0,811449945	0,826666339	0,893814849

5. Másod (és magasabb) rendű lineáris differenciálegyenletek

Egy másodrendű lineáris differenciálegyenlet általános alakja:

$$y''(x) + p_1(x)y'(x) + p_2(x)y(x) = q(x) \quad (74)$$

ahol a $p_1(x)$, $p_2(x)$, $q(x)$ együtthatók értelmezve vannak egy közös intervallumon és $y(x)$ az ismeretlen függvény.

5.1. Definíció.

Ha a fenti egyenletben $q(x) \equiv 0$, akkor a differenciálegyenlet *homogén lineáris*. Ellenkező esetben *inhomogén*.

5.1. Tétel.

Ha $y_1(x)$ és $y_2(x)$ megoldása a

$$y''(x) + p_1(x)y'(x) + p_2(x)y(x) = 0 \quad (75)$$

másodrendű homogén lineáris differenciálegyenletnek, akkor tetszőleges c_1 és c_2 konstansokkal az

$$y(x) = c_1y_1(x) + c_2y_2(x)$$

függvény is megoldás.

5.2. Tétel.

Ha $y_1(x)y_2'(x) - y_1'(x)y_2(x) \neq 0$, akkor a másodrendű homogén lineáris differenciálegyenlet bármely megoldása előáll $y(x) = c_1y_1(x) + c_2y_2(x)$ alakban alkalmas c_1 és c_2 konstansokkal.

5.2. Definíció.

Ha $y_1(x)y_2'(x) - y_1'(x)y_2(x) \neq 0$, akkor azt mondjuk, hogy az $y_1(x)$ és $y_2(x)$ függvények függetlenek.

5.1. Megjegyzés.

A fenti tételekből következik, hogy egy másodrendű homogén lineáris differenciálegyenlet általános megoldásának meghatározásához elég megtalálni két független megoldását.

5.3. Tétel.

Ha $y_1(x)$ és $y_2(x)$ függvények független megoldásai a másodrendű homogén lineáris differenciálegyenletnek és $y_0(x)$ egy megoldása az

$$y''(x) + p_1(x)y'(x) + p_2(x)y(x) = q(x)$$

másodrendű inhomogén lineáris differenciálegyenletnek, akkor ennek az egyenletnek minden megoldása előállítható

$$y(x) = c_1y_1(x) + c_2y_2(x) + y_0(x)$$

alakban.

Következmény:

Egy másodrendű inhomogén lineáris differenciálegyenlet általános megoldásához meg kell találni a homogén rész két független megoldását $y_1(x)$ és $y_2(x)$ -t, és az inhomogén egyenlet egy (úgynevezett partikuláris) megoldását $y_0(x)$ -t.

5.4. Tétel. Egy állandó variálása

Ha $y_1(x)$ megoldása a másodrendű homogén lineáris differenciálegyenletnek, akkor

$$y_2(x) = y_1(x) \int \frac{1}{y_1^2(x)} e^{-\int p_1(x) dx} dx$$

függvény az $y_1(x)$ függvénytől független másik megoldása.

5.5. Tétel. Két állandó variálása

Ha $y_1(x)$ és $y_2(x)$ függvények független megoldásai a másodrendű homogén lineáris differenciálegyenletnek, akkor az

$$y_0(x) = \left(\int \frac{-y_2(x)q(x)}{W(x)} dx \right) y_1(x) + \left(\int \frac{y_1(x)q(x)}{W(x)} dx \right) y_2(x)$$

függvény megoldása az inhomogén lineáris differenciálegyenletnek, ahol

$$W(x) = y_1(x)y_2'(x) - y_1'(x)y_2(x) \neq 0$$

5.2. Megjegyzés.

A fenti két tételből következik, hogy ha tudjuk a homogén rész egy megoldását, akkor elő tudjuk állítani az inhomogén egyenlet általános megoldását. Sajnos általános esetben nincs univerzális módszer a homogén rész egy megoldásának megtalálására. Speciális esetekben tudjuk csak meghatározni.

A következő példán megmutatjuk hogyan lehet egy másodrendű inhomogén lineáris differenciálegyenlet általános megoldását megtalálni a homogén rész egy megoldásának ismeretében. Mivel nem fogjuk felhasználni a fenti két tételben szereplő képleteket, az eljárásból majd az is kiderül, miért nevezik a módszert egy, illetve két állandó variálásának.

5.1. Példa.

Oldjuk meg az

$$y'' + \frac{x}{1-x}y' - \frac{1}{1-x}y = 1 - x$$

differenciálegyenletet, ha tudjuk, hogy az $y_1 = e^x$ függvény kielégíti a homogén részét.

Megoldás: Először is könnyű ellenőrizni, hogy az $y_1 = e^x$ függvény valóban kielégíti az egyenlet homogén részét.

Ezután egy állandó variálásával meghatározzuk a homogén rész egy másik megoldását. Keresük az y_2 függvényt $y_2 = c(x)y_1$ alakban. Tehát

$$\begin{aligned} y_2 &= c(x)e^x \\ y_2' &= c'(x)e^x + c(x)e^x \\ y_2'' &= c''(x)e^x + 2c'(x)e^x + c(x)e^x \end{aligned}$$

Behelyettesítve a differenciálegyenlet homogén részébe

$$c''(x)e^x + 2c'(x)e^x + c(x)e^x + \frac{x}{1-x} [c'(x)e^x + c(x)e^x] - \frac{1}{1-x} c(x)e^x = 0$$

Rendezve

$$c'' + \left(2 + \frac{x}{1-x}\right) c' = 0$$

adódik. Ebből

$$\frac{c''}{c'} = -1 + \frac{1}{x-1}$$

Integrálva és az integrációs konstanszt zérusnak választva

$$\ln c' = -x + \ln(x-1)$$

azaz

$$c' = (x-1)e^{-x}$$

Újból integrálva

$$c = -xe^{-x}$$

Tehát

$$y_2 = c(x)e^x = -x$$

Következésképpen a differenciálegyenlet homogén részének általános megoldása:

$$y_h = c_1 e^x + c_2 x.$$

Most már ennek ismeretében a két állandó c_1 és c_2 variálásával - azaz helyükre függvényeket írva megkeressük az inhomogén egyenlet egy megoldását. Legyen

$$y_0 = c_1(x)y_1 + c_2(x)y_2 = c_1(x)e^x + c_2(x)x$$

$$y_0' = c_1'(x)e^x + c_1(x)e^x + c_2'(x)x + c_2(x)$$

A keresett $c_1(x)$ és $c_2(x)$ függvényekre még egy feltételt írhatunk elő. Legyen

$$(*) \quad c_1'(x)e^x + c_2'(x)x = 0$$

Ekkor a második derivált

$$y_0'' = c_1'(x)e^x + c_1(x)e^x + c_2'(x).$$

Behelyettesítve az inhomogén egyenletbe:

$$[c_1'(x)e^x + c_1(x)e^x + c_2'(x)] + \frac{x}{1-x} [c_1(x)e^x + c_2(x)] - \frac{1}{1-x} [c_1(x)e^x + c_2(x)x] = 1 - x$$

Rendezve adódik, hogy

$$(**) \quad c_1'(x)e^x + c_2'(x) = 1 - x$$

A (*) és (**) egyenleteket megoldjuk, és kapjuk, hogy

$$c_1'(x) = -xe^{-x} \quad c_2'(x) = 1$$

Integrálva

$$c_1(x) = (x + 1)e^{-x} \quad c_2(x) = x$$

Tehát

$$y_0 = x + 1 + x^2$$

A példabeli differenciálegyenlet általános megoldása:

$$y = c_1e^x + c_2x + x^2 + x + 1.$$

5.1. Állandó együtthatós eset

Most azzal a speciális esettel foglalkozunk, amikor a

$$y''(x) + p_1(x)y'(x) + p_2(x)y(x) = q(x) \quad (76)$$

differenciálegyenletben a $p_1(x)$, $p_2(x)$, együtthatók helyett p_1 és p_2 konstansok állnak:

$$y''(x) + p_1y'(x) + p_2y(x) = q(x) \quad (77)$$

A differenciálegyenlet általános megoldása most is

$$y(x) = c_1y_1(x) + c_2y_2(x) + y_0(x),$$

de a homogén rész két független $y_1(x)$ illetve $y_2(x)$ megoldásának meghatározása lényegesen egyszerűbb. Keressük ugyanis a homogén rész megoldását $y = e^{\lambda x}$ alakban. Ezt a függvényt, és deriváltjait behelyettesítve a homogén egyenletbe adódik, hogy

$$\lambda^2 e^{\lambda x} + p_1 \lambda e^{\lambda x} + p_2 e^{\lambda x} = 0.$$

Mivel $e^{\lambda x} \neq 0$, egyszerűsíthetjük az egyenletet:

$$\lambda^2 + p_1 \lambda + p_2 = 0.$$

Ez egy egyszerű másodfokú egyenlet. Három esetet tekintünk.

- I. $p_1^2 - 4p_2 > 0$ két különböző valós gyök. Legyenek λ_1 és λ_2 . Ekkor az

$$y_1 = e^{\lambda_1 x} \quad y_2 = e^{\lambda_2 x}$$

két független megoldása a differenciálegyenletnek.

- II. $p_1^2 - 4p_2 = 0$ két egybeeső valós gyök $\lambda = -\frac{p_1}{2}$. Ebben az esetben könnyen igazolható, hogy

$$y_1 = e^{\lambda x} \quad y_2 = xe^{\lambda x}$$

két független megoldása a differenciálegyenletnek.

- $p_1^2 - 4p_2 < 0$ két különböző komplex gyök, amelyek egymás komplex konjugáltjai. Jelöljük őket $\alpha + i\beta$ illetve $\alpha - i\beta$ -val. Ekkor a homogén lineáris differenciálegyenlet két független megoldása:

$$y_1 = e^{\alpha x} \cos \beta x \quad \text{és} \quad y_2 = e^{\alpha x} \sin \beta x.$$

5.2. Magasabb rendű állandó együtthatós lineáris differenciálegyenletek

Az n -edrendű állandó együtthatós

$$y^{(n)}(x) + p_1 y^{(n-1)}(x) + \dots + p_n y(x) = q(x) \quad (78)$$

lineáris differenciálegyenlet megoldása a másodrendűéhez hasonlóan történik.

A homogén rész megoldását alkotó független megoldásfüggvényeket most is $e^{\lambda x}$ alakban keressük. Ebből adódik a

$$\lambda^n + p_1 \lambda^{n-1} + \dots + p_n = 0$$

egyenlet. Ennek ha valamely λ valós szám egyszeres gyöke, akkor az $y = e^{\lambda x}$ megoldása a homogén egyenletnek.

Ha a λ valós szám r -szeres gyök, akkor az

$$e^{\lambda x}, \quad x e^{\lambda x}, \quad \dots, \quad x^{r-1} e^{\lambda x}$$

független megoldásai a homogén egyenletnek.

Ha pedig λ komplex gyök, akkor a komplex konjugáltja is gyök, ezért az ehhez tartozó megoldások most is

$$e^{\alpha x} \cos \beta x \quad \text{és} \quad e^{\alpha x} \sin \beta x.$$

Az inhomogén egyenlet egy partikuláris megoldásának keresése ugyan úgy történhet, mint a másodrendű esetben, csak most egy n egyenletből álló egyenletrendszert kell megoldani a c_i függvények deriváltjaira.

5.3. Másodrendű lineáris differenciálegyenletek integrálása sorok segítségével

Az

$$y'' + p(x)y' + q(x)y = 0$$

egyenlet megoldását, vagy annak közelítését keressük.

Fejtsük hatványsorba az egyenletben szereplő együtthatókat:

$$p(x) = \sum_{i=0}^{\infty} a_i x^i \quad \text{és} \quad q(x) = \sum_{i=0}^{\infty} b_i x^i$$

A megoldást is hatványsor alakban keressük:

$$y(x) = \sum_{i=0}^{\infty} c_i x^i$$

Ebből

$$y'(x) = \sum_{i=0}^{\infty} i \cdot c_i x^{i-1}$$

$$y''(x) = \sum_{i=1}^{\infty} i \cdot (i-1) \cdot c_i x^{i-2}$$

Ahhoz, hogy két független megoldást kapjunk, először a $c_0 = 1, c_1 = 0$ választással meghatározzuk y_1 függvényt, majd $c_0 = 0, c_1 = 1$ választással y_2 függvényt. A c_i együtthatók meghatározása úgy történik, hogy a hatványsorokat behelyettesítjük a differenciálegyenletbe és az azonos hatványok együtthatóit zérussal tesszük egyenlővé.

$$\begin{array}{l|l} x^0 & 2 \cdot 1 \cdot c_2 + a_0 c_1 + b_0 c_0 = 0 \\ x^1 & 3 \cdot 2 \cdot c_3 + 2a_0 c_2 + a_1 c_1 + b_0 c_1 + b_1 c_0 = 0 \\ x^2 & 4 \cdot 3 \cdot c_4 + 3a_0 c_3 + 2a_1 c_2 + a_2 c_1 + b_0 c_2 + b_1 c_1 + b_2 c_0 = 0 \\ \vdots & | \end{array}$$

5.4. Véges-elem módszer

A véges-elem módszer hatékonysága miatt nagyon gyakran használt módszer mechanikai, statikai problémák megoldására. Éppen ezért ma már könyvtárnyi irodalma van. Itt most az egyszerűség kedvéért egy példán keresztül próbáljuk bemutatni a lényegét.

Keressünk egy $y(x)$ függvényt, amely értelmezve van a $[0; 1]$ intervallumon és kielégíti a

$$-y'' + y = x \tag{79}$$

differenciálegyenletet az $y(0) = y(1) = 0$ peremfeltételekkel.

5.3. Megjegyzés.

Itt persze ki lehet találni a pontos megoldást: $y(x) = x - \frac{sh(x)}{sh(1)}$, de az esetek többségében olyankor használják a végeselem módszert, amikor a megoldás nem állítható elő explicite.

Megoldás:

Szorozzuk meg a (79) differenciálegyenletet egy $w(x)$ súlyfüggvénnyel, amely eleget tesz a következő feltételeknek:

A $w(x)$ súlyfüggvény deriváltjának négyzete integrálható a $[0; 1]$ intervallumon, és $w(0) = w(1) = 0$.

A (79) egyenletből a $w(x)$ súlyfüggvénnyel való szorzás után integrálással kapjuk, hogy

$$\int_0^1 [-y''(x) + y(x)]w(x)dx = \int_0^1 xw(x)dx. \tag{80}$$

A bal oldali integrál első tagját parciálisan integrálva

$$\int_0^1 -y''w dx = \int_0^1 y'w' dx - [y'w]_0^1$$

A $w(0) = w(1) = 0$ feltétel miatt $[y'w]_0^1 = 0$, ezért az

$$\int_0^1 -y''w dx = \int_0^1 y'w' dx$$

egyenletet kapjuk. Ezt felhasználva a (80) egyenlet így alakul:

$$\int_0^1 (y'w' + yw)dx = \int_0^1 xw dx \quad (81)$$

Ennek az egyenletnek fogjuk keresni közelítő megoldását az úgynevezett **Galjorkin-approximációval**.
Legyen

$$\Phi_1, \Phi_2, \Phi_3, \dots$$

a $[0; 1]$ intervallumon folytonos függvényeknek egy olyan sorozata, amelyre igaz, hogy bármely, a feltételeknek eleget tevő $w(x)$ súlyfüggvény, előáll a $\Phi_i(x)$ függvények segítségével

$$w(x) = \sum_{i=1}^{\infty} \beta_i \Phi_i(x)$$

alakban. Legyen

$$w_N(x) = \sum_{i=1}^N \beta_i \Phi_i(x)$$

a $w(x)$ súlyfüggvény egy approximációja. Feltesszük, hogy

$$\lim_{N \rightarrow \infty} w_N = w$$

A (81) differenciálegyenlet $y(x)$ megoldásának pedig legyen

$$y_N(x) = \sum_{i=1}^N \alpha_i \Phi_i(x)$$

egy alkalmas közelítése, amelyről azt reméljük, hogy

$$\lim_{N \rightarrow \infty} y_N = y(x).$$

Helyettesítsük be ezeket a (81) differenciálegyenletbe.

$$\int_0^1 \left[\sum_{i=1}^N \alpha_i \Phi_i' \sum_{k=1}^N \beta_k \Phi_k' - \sum_{i=1}^N \alpha_i \Phi_i \sum_{k=1}^N \beta_k \Phi_k \right] dx = \int_0^1 x \cdot \left[\sum_{k=1}^N \beta_k \Phi_k \right] dx$$

Ezt az egyenletet átrendezve kapjuk, hogy

$$\sum_{i=1}^N \beta_i \left\{ \sum_{k=1}^N \left[\int_0^1 (\Phi_i' \Phi_k' + \Phi_i \Phi_k) dx \right] \alpha_k - \int_0^1 x \Phi_i(x) dx \right\} = 0.$$

Ez az egyenlet a beta konstansok tetszőleges megválasztása mellett csak úgy állhat fenn, ha a kapcsos zárójelen belüli rész zérus. Jelölje

$$K_{i,k} = \int_0^1 (\Phi_i' \Phi_k' + \Phi_i \Phi_k) dx$$

és

$$F_i = \int_0^1 x \Phi_i(x) dx$$

Ekkor egyenletünk az alábbi lesz

$$\sum_{i=1}^N (K_{i,k} \alpha_k - F_i) = 0.$$

Ugyan ez más alakban:

$$\mathbf{K} \vec{\alpha} = \vec{F},$$

ahol

$$\vec{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{bmatrix} \quad \vec{F} = \begin{bmatrix} F_1 \\ \vdots \\ F_N \end{bmatrix}$$

A \mathbf{K} mátrix szokásos elnevezése: **merevségi mátrix**, az \vec{F} vektor pedig a **teher vektor**. Mivel a Φ_i függvények függetlenek, ezért megválaszthatók úgy, hogy a \mathbf{K} mátrix invertálható legyen, és ekkor

$$\vec{\alpha} = \mathbf{K}^{-1} \vec{F}.$$

Ezzel az y_N közelítő megoldást meghatároztuk.

Most már csak az a kérdés, hogyan válasszuk meg a Φ_i függvényeket.

a) Ha a Φ_i függvényeket a trigonometrikus rendszer szerint választjuk, akkor a fenti módszert *Fourier-módszernek* nevezzük.

b) Ha a Φ_i függvényeket $\Phi_i(x) = x^i(1-x)$ választjuk, akkor a fenti módszert *Ritz-módszernek* nevezzük.

c) Ha a Φ_i függvényeknek a (1.3) definícióban meghatározott *sátorfüggvényeket* választjuk, akkor a \mathbf{K} mátrix "sávmátrix" lesz, azaz csak a főátlóban és közvetlen alatta, fölötté lesznek zérustól különböző elemek. Ebben az esetben a \mathbf{K} merevségi mátrix és \vec{F} teher vektor könnyen számolható. A megoldás közelítése pedig szakaszonként lineáris.

5.5. Másodrendű differenciálegyenletek mechanikai alkalmazásai

Differenciálegyenletek mechanikai alkalmazására már a bevezetőben láttunk két példát (4.1) és (4.2) példa. Most további két példa következik.

5.2. Példa.

Gerenda lehajlása.

Egy L hosszú gerenda vízszintesen egyik végén van befogva, másik végén P teher. Mekkora lehajlás jön létre a rúd végén?

ide kell ábra

Megoldás:

A befogástól x távolságban létrejövő forgatónyomaték:

$$M = P(L - x)$$

A fizikai törvények szerint a keletkező görbület arányos a létrehozó forgatónyomatékkal. Az arányossági tényező függ az anyagi minőségtől (E) és a rúd keresztmetszetének alakjától és méreteitől (I).

$$\frac{y''}{[1 + (y')^2]^{3/2}} = \frac{P(L - x)}{EI}$$

Ezt az egyenletet nem tudjuk explicite megoldani, de ha feltesszük, hogy csak kis deformáció jöhet létre, azaz $y' \approx 0$, akkor közelítőleg

$$y'' = \frac{P(L - x)}{EI}$$

Ezt kétszer integrálva kapjuk, hogy

$$y = \frac{P}{EI} \cdot \frac{(L - x)^3}{6} + c_1x + c_2$$

Használjuk fel, hogy a rögzített végnél $x = 0$ esetén $y = 0$ és $y' = 0$, ezért

$$c_1 = \frac{P}{EI} \cdot \frac{L^2}{2}$$

és

$$c_2 = -\frac{P}{EI} \cdot \frac{L^3}{6}$$

ezért

$$y = \frac{P}{EI} \cdot \left(L\frac{x^2}{2} - \frac{x^3}{6} \right).$$

Tehát a lehajlás a rúd végén

$$h = \frac{P}{EI} \cdot \frac{L^3}{3}.$$

5.3. Példa.

Rezgés, kényszerrezgés

- Tegyük fel, hogy egy m tömegű testre egyedül egy rugóerő hat. Szeretnénk meghatározni, hogy mozog a test. Felírva Newton második törvényét:

$$m\ddot{x} = -kx,$$

ahol x az elmozdulás, annak idő szerinti második deriváltja \ddot{x} a gyorsulás és k a rugóállandó. Mivel k és m pozitív állandók, vezessük be a $\frac{k}{m} = \omega^2$ jelölést. Megoldandó tehát a

$$\ddot{x} + \omega^2 x = 0$$

állandóegyütthetős másodrendű lineáris homogén differenciálegyenlet $x(0) = x_0$ és $\dot{x} = v_0$ kezdeti feltételekkel. A differenciálegyenlet karakterisztikus egyenlete:

$$\lambda^2 + \omega^2 = 0,$$

aminek megoldása $\lambda = \pm \omega i$. Ezért az általános megoldás:

$$x(t) = c_1 \cdot \cos \omega t + c_2 \cdot \sin \omega t.$$

A c_1 és c_2 konstansok a kezdeti feltételből határozhatók meg. Például, ha a mozgást az origóból indítjuk $x(0) = 0$ $v_0 = \dot{x}(0)$ kezdősebességgel, akkor

$$x(t) = \frac{v_0}{\omega} \cdot \sin \omega t$$

szinuszos rezgőmozgást kapjuk. (Csillapítatlan rezgés.)

- Most tegyük fel, hogy van közegellenállás, ami a sebességgel arányosan akadályozza a mozgást. Ekkor a differenciálegyenlet:

$$\ddot{x} + \beta \dot{x} + \omega^2 x = 0.$$

A karakterisztikus egyenlet:

$$\lambda^2 + \beta \cdot \lambda + \omega^2 = 0.$$

Ennek gyökei:

$$\lambda = \frac{-\beta \pm \sqrt{\beta^2 - 4\omega^2}}{2}$$

Két eset van:

a) β , a közegellenállás olyan nagy, hogy $\beta^2 - 4\omega^2 \geq 0$. Ekkor a karakterisztikus egyenletnek valós gyökei vannak és a megoldás:

$$x(t) = c_1 \cdot e^{\lambda_1 t} + c_2 \cdot e^{\lambda_2 t}.$$

A kezdeti feltételtől függően a nyugalmi helyzetéből kitérített test exponenciálisan (periódikus mozgás nélkül = aperiódikusan) visszatér a nyugalmi helyzetéhez.

b) A közegellenállás olyan kicsi, hogy $\beta^2 - 4\omega^2 < 0$. Ebben az esetben a karakterisztikus egyenlet gyökei komplex számok. Jelölje $-\frac{\beta}{2} \pm \hat{\omega}i$. A megoldás:

$$x(t) = c_1 e^{-\frac{\beta}{2} t} \cos \hat{\omega} t + c_2 e^{-\frac{\beta}{2} t} \sin \hat{\omega} t.$$

Ekkor tehát olyan csillapodó rezgőmozgás jön létre, ahol az amplitúdó exponenciálisan lecseng.

- Kényszerrezgés, periodikus kényszer.
Most hadd hasonlítsunk a rendszerre periodikus kényszerű ω_k kényszerfrekvenciával. A differenciálegyenlet:

$$\ddot{x} + \beta\dot{x} + \omega^2 x = A \sin \omega_k t.$$

Ha ω_k nem gyöke a karakterisztikus egyenletnek, akkor az általános megoldás

$$x(t) = c_1 \cdot \cos \omega_k t + c_2 \cdot \sin \omega_k t$$

azaz a rezgés a kényszerfrekvencián történik.

Ha ω_k gyöke a karakterisztikus egyenletnek (rezonancia), akkor a megoldás

$$x(t) = c_1 \cdot t \cdot \cos \omega_k t + c_2 \cdot t \cdot \sin \omega_k t$$

alakú. Az amplitúdó lineárisan megnő. Ezt nevezik rezonanciakatasztrófának.

6. Valószínűségszámítás alapjai

A valószínűségszámítás véletlentől függő tömegjelenségekkel kapcsolatos eseményekkel foglalkozik. A véletlen tömegjelenségek törvényszerűségeit vizsgálja.

Az események ugyanazon kísérlet kimenetelével függnek össze. Pl. Kísérlet: kockadobás. Esemény: kettést dobunk, páratlan számot dobunk, stb. Pl. Egy fődém gerendát G teherrel terhelünk Esemény: adott helyen x lehajlást mérünk.

Az események matematikai tulajdonságai: Két esemény A és B azonos, ha a kísérlet bármely kimenetele esetén vagy mindkettő bekövetkezik vagy egyik sem. $A = B$. Valamilyen A esemény be nem következése is esemény. Ennek jelölése: \bar{A} .

Az eseményeken műveleteket értelmezzük.

6.1. Definíció.

Összeadás:

A és B események összege egy olyan esemény, ami akkor következik be, ha az A és B események közül legalább az egyik bekövetkezik. Jelölése: $A + B$.

6.2. Definíció.

Szorzás:

A és B események szorzata egy olyan esemény, ami akkor következik be, ha az A és B események közül mindkettő bekövetkezik. Jelölése: AB .

Van két kitüntetett esemény: Az egyik a biztos esemény, amely a kísérlet bármely kimenetele esetén bekövetkezik. jelölése I . (Pl. kockadobás esetén 7-nél kisebbet dobunk.)

A másik a lehetetlen esemény, amely a kísérlet bármely kimenetele esetén nem következik be.

Jelölése: \emptyset . (Pl. kockadobás esetén tízest dobunk)

Műveleti tulajdonságok:

$$\begin{array}{ll} A + B = B + A & AB = BA \\ (A + B) + C = A + (B + C) & (AB)C = A(BC) \\ A + A = A & AA = A \\ A + \emptyset = A & A\emptyset = \emptyset \\ A + I = I & AI = A \\ A + \bar{A} = I & A\bar{A} = \emptyset \\ (A + B)C = AC + BC & AB + C = (A + C)(B + C) \end{array}$$

Relatív gyakoriság Tekintsünk egy olyan kísérletet, amelynél a figyelembe vett körülmények a kísérlet eredményét nem határozzák meg egyértelműen, hanem többféle kimenetelt engednek meg. Legyen az A esemény ezeknek a lehetőségeknek az egyike. Hajtsuk végre a kísérletet többször, azonos körülmények között. Ekkor az A esemény a kísérletek egy részében bekövetkezik, más részében nem. Az utóbbi esetben az A esemény ellentettje következik be, amelynek jelölése: \bar{A} . Ha n kísérletből az A esemény pontosan k -szor következett be, akkor k -t az esemény gyakoriságának, $\frac{k}{n}$ -et pedig az esemény *relatív gyakoriságának* nevezzük az illető kísérletsorozatban.

Egy véletlen esemény relatív gyakorisága különböző kísérletsorozatban általában nem állandó, de gyakran azt tapasztaljuk, hogy egy bizonyos érték körül ingadozik. Pl. dobjunk fel egy pénzérmét 100-szor és számoljuk meg hányszor kaptunk fejet. Többször elvégezve a 100 dobásból álló kísérletsorozatot azt tapasztaljuk, hogy a fejek számának relatív gyakorisága az $1/2$ körül ingadozik.

A relatív gyakoriság alapvető tulajdonságai:

- Egy esemény relatív gyakorisága nem negatív szám és legfeljebb 1.
- A biztos esemény relatív gyakorisága 1.
- Ha A és B ugyanazon kísérlethez tartozó egymást kizáró események (azaz egyszerre nem következhetnek be, vagyis $AB = \emptyset$), akkor a két esemény összegének a relatív gyakorisága megegyezik az események relatív gyakoriságának összegével.

6.1. A valószínűség fogalma

6.3. Definíció.

A valószínűség eseményekhez rendelt szám (eseményeken értelmezett függvény). Az A esemény valószínűségét $P(A)$ -val jelöljük. A valószínűség a relatív gyakoriság mintájára a következő tulajdonságú:

- $P(A) \geq 0$
- $P(I) = 1$
- Ha $AB = \emptyset$, akkor $P(A + B) = P(A) + P(B)$.

Két fontos következmény:

- A lehetetlen esemény valószínűsége zérus: $P(\emptyset) = 0$.
- $P(A + B) = P(A) + P(B) - P(AB)$.

6.2. Klasszikus valószínűségi mező

Ha egy kísérlet olyan, hogy megfogalmazható benne véges sok (n darab) egyformán valószínű esemény, amelyek közül néhánynak az összegével az összes többi esemény kifejezhető, akkor ezt az n darab eseményt *elemi eseménynek* nevezzük. Ezek összegeként előállítható eseményeket összetett eseménynek hívjuk. Az elemi események valószínűsége $\frac{1}{n}$. Ha egy összetett esemény k darab elemi eseményből tevődik össze, akkor annak valószínűsége $\frac{k}{n}$. (Ilyenkor szokták azt mondani, hogy az esemény valószínűsége a kedvező esetek száma osztva az összes esetek számával.)

6.3. Feltételes valószínűség

6.4. Definíció.

Ha $P(B) > 0$, akkor az A eseménynek a B eseményre vonatkozó feltételes valószínűségét a $P(A|B) = \frac{P(AB)}{P(B)}$ képlettel definiáljuk.

6.1. Példa.

Legyen A esemény az, hogy a kockával 6-ost dobunk, B pedig az az esemény, hogy a dobás eredménye páros szám.

Mennyi a valószínűsége, hogy a dobás eredménye 6-os, feltéve, hogy tudjuk a dobás páros számot eredményezett?

Válasz:

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$$

6.5. Definíció.

Függetlenség:

Az A és B események függetlenek, ha az A eseménynek a B eseményre vonatkozó feltételes valószínűsége megegyezik az A esemény feltétel nélküli valószínűségével, azaz $P(A|B) = P(A)$.

Ekkor

$$\frac{P(AB)}{P(B)} = P(A),$$

vagyis

$$P(AB) = P(A)P(B).$$

Ez utóbbi képlet mutatja, hogy a függetlenség szimmetrikus, azaz ha A független a B eseménytől, akkor B is független A -tól.

A feltételes valószínűségekre vonatkozik a következő fontos tétel:

6.1. Tétel.

Bayes tétel (1)

Ha $P(B) \neq 0$, akkor

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

6.6. Definíció.

Teljes eseményrendszer

Az A_1, A_2, \dots, A_n események teljes eseményrendszert alkotnak, ha páronként kizáró események és összegük a biztos esemény, azaz $A_i A_j = \emptyset$ ha $i \neq j$ és $\sum_{i=1}^n A_i = I$.

6.2. Tétel.

Teljes valószínűség tétele

Ha az A_1, A_2, \dots, A_n események teljes eseményrendszert alkotnak és $P(A_i) \neq 0$, akkor tetszőleges B eseményre

$$P(B) = \sum_{i=1}^n P(B|A_i)P(A_i).$$

6.2. Példa.

Tegyük fel, hogy a BME-n tanuló lányok 80%-a, a fiúk 65%-a jól beszél angolul. Mi a valószínűsége, hogy ha egy véletlenül szembejövő hallgatót megszólítottunk, akkor tudunk vele angolul társalogni? Tudjuk azt, hogy a hallgatók 40%-a lány, 60 %-a fiú.

Megoldás: Jelölje B azt az eseményt, hogy egy hallgató jól beszél angolul, és legyen A_1 az az esemény, hogy egy véletlenül meglátott hallgató lány, A_2 pedig az az esemény, hogy fiú. Látjuk, hogy A_1 és A_2 teljes eseményrendszert alkot, mert egymást kizáró események és összegük a biztos esemény.

A feltétel szerint

$$P(B|A_1) = 0,8 \quad P(B|A_2) = 0,65 \quad P(A_1) = 0,4 \quad P(A_2) = 0,6$$

Tehát

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) = 0,8 \cdot 0,4 + 0,65 \cdot 0,6 = 0,71$$

6.3. Tétel.

Bayes tétel (2)

Ha az A_1, A_2, \dots, A_n események teljes eseményrendszert alkotnak és $P(B) \neq 0$, akkor

$$P(A_k|B) = \frac{P(B|A_k)P(A_k)}{\sum_{i=1}^n P(B|A_i)P(A_i)}.$$

6.3. Példa.

7. Diszkrét valószínűségi változó

Valószínűségi változónak nevezzünk egy, az eseménytíren értelmezett függvényt, azaz az eseményekhez rendelünk számokat.

Pl. Legyen a kísérlet az, hogy két pénzdarabot feldobunk. Legyen a valószínűségi változó a feldobások száma. Ekkor tehát a valószínűségi változó lehetséges értékei 0, 1, vagy 2. Az ezekhez tartozó valószínűség értékek: $1/4, 1/2, 1/4$.

Tegyük fel, hogy egy valószínűségi változónak véges sok vagy megszámlálhatóan végtelen sok lehetséges értéke van. Az ilyen valószínűségi változót *diszkrét valószínűségi változónak* nevezzük. Ekkor a valószínűségi változót megadhatjuk úgy, hogy felsoroljuk a lehetséges értékeit:

$$x_1, x_2, \dots, x_n, \dots$$

és a hozzátartozó valószínűség-eloszlást:

$$p_1, p_2, \dots, p_n, \dots$$

Ha ξ jelöl egy valószínűségi változót, amelynek lehetséges értékei $x_1, x_2, \dots, x_n, \dots$, akkor a valószínűségeloszlása: $P(\xi = x_i) = p_i$.

7.1. Várható érték, szórás

Egy valószínűségi változó fontos jellemzője a *várható érték*, ami a lehetséges értékeknek a valószínűségekkel súlyozott középértéke. Diszkrét esetben:

7.1. Definíció.

$$E[\xi] = \sum_i p_i x_i$$

7.2. Definíció.

Egy valószínűségi változó *szórásnégyzete* a várható értéktől való eltérés négyzetének a várható értéke, azaz

$$D^2[\xi] = E[(\xi - E[\xi])^2]. \quad (82)$$

A szórás a szórásnégyzet pozitív négyzetgyöke: $D[\xi]$. Kiszámítása:

$$D^2[\xi] = \sum_i (x_i - E[\xi])^2 p_i,$$

vagy

$$\begin{aligned} D^2[\xi] &= \sum_i (x_i - E[\xi])^2 p_i = \sum_i (x_i^2 - 2x_i E[\xi] + E^2[\xi]) p_i = \\ &= \sum_i x_i^2 p_i - 2E[\xi] \sum_i x_i p_i + E^2[\xi] \sum_i p_i = \sum_i x_i^2 p_i - 2E^2[\xi] + E^2[\xi] = \\ &= \sum_i x_i^2 p_i - E^2[\xi] = E[\xi^2] - E^2[\xi] \end{aligned}$$

7.2. Nevezetes diszkrét eloszlások

(i) A legegyszerűbb nevezetes diszkrét valószínűségi változó az *egyszerű alternatíva*. Olyankor lép fel, amikor csak az érdekel bennünket, hogy egy esemény bekövetkezik-e vagy sem. Legyen a valószínűségi változó :

$$\xi = \begin{cases} 0 & \text{ha a megfigyelt esemény nem következik be} \\ 1 & \text{ha a megfigyelt esemény bekövetkezik} \end{cases}$$

A hozzátartozó valószínűségeloszlás:

$$P(\xi = 0) = q = 1 - p \quad P(\xi = 1) = p$$

Ennek várható értéke és szórása:

$$E[\xi] = p \quad D[\xi] = pq$$

(ii) *Binomiális eloszlás* vagy Bernoulli eloszlású valószínűségi változó esetén egy alternatívát n -szer egymástól függetlenül megismételünk és azt figyeljük, hogy hányszor következett be a kérdéses esemény. A binomiális eloszlású valószínűségi változó előállítható az n darab független alternatíva összegeként. Ha η jelöli ezt a valószínűségi változót, akkor

$$\eta = \xi_1 + \xi_2 + \dots + \xi_n$$

Tehát η lehetséges értékei: $0, 1, 2, \dots, n$. Valószínűségeloszlása:

$$P(\eta = k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} p^k q^{n-k}$$

Várható értéke és szórása:

$$E[\eta] = np \quad D[\eta] = \sqrt{npq}$$

Tipikus alkalmazása a visszatevéses mintavétel.

(iii) *Hipergeometriai eloszlás*. Tipikus alkalmazása a visszatevés nélküli mintavétel. Tegyük fel, hogy van N darab termék, amelyből M darab hibás. Kiveszünk egy n elemű mintát. Legyen a valószínűségi változó értéke a mintában talált selejtek száma. A valószínűségi változó lehetséges értékei most is $0, 1, 2, \dots, n$. Valószínűségeloszlása:

$$P(\zeta = k) = \frac{\binom{N}{k} \binom{N-M}{n-k}}{\binom{N}{n}}$$

Várható értéke és szórása:

$$E[\zeta] = n \frac{M}{N} \quad D[\zeta] = \sqrt{n \frac{M}{N} \left(1 - \frac{M}{N}\right) \left(1 - \frac{n-1}{N-1}\right)}$$

(iv) *Geometriai eloszlás.* Most egy alternatívát addig ismételünk, amíg először bekövetkezik a megfigyelni kívánt esemény. A valószínűségi változó lehetséges értékei: $1, 2, \dots$. Valószínűségeloszlása:

$$P(\gamma = k) = pq^{k-1}.$$

Várható értéke és szórása:

$$E[\gamma] = \frac{1}{p} \quad D[\gamma] = \frac{\sqrt{q}}{p}.$$

(v) *Poisson eloszlás.* A Poisson eloszlást a binomiális eloszlás határeseteként kapjuk, ha $n \rightarrow \infty$ és $p \rightarrow 0$, úgy hogy $np \rightarrow \lambda > 0$. A valószínűségi változó lehetséges értékei tehát $0, 1, 2, \dots$. Valószínűségeloszlása:

$$P(\theta = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Várható értéke és szórása:

$$E[\theta] = \lambda \quad D[\theta] = \sqrt{\lambda}.$$

8. Folytonos valószínűségi változó

Egy ξ diszkrét valószínűségi változó esetén fel lehet sorolni a valószínűségi változó lehetséges értékeit és megadni hozzá a $P(\xi = x_i)$ valószínűség eloszlást. Ha a valószínűségi változó a lehetséges értékeit egy véges, vagy végtelen intervallumon veszi fel, akkor a valószínűség eloszlás definiálására az

$$F(x) = P(\xi < x)$$

eloszlás függvényt használjuk.

Az eloszlás függvény értelmezéséből következnek az alábbi alapvető tulajdonságai:

- $$0 \leq F(x) \leq 1$$

- $$\lim_{x \rightarrow -\infty} F(x) = 0$$

- $$\lim_{x \rightarrow +\infty} F(x) = 1$$

- $F(x)$ monoton növekvő.

A továbbiakban feltesszük, hogy az $F(x)$ eloszlásfüggvény nem csak folytonos, de differenciálható is.

8.1. Definíció.

Az $F(x)$ eloszlásfüggvény $F'(x) = f(x)$ deriváltját *sűrűségfüggvénynek* nevezzük. A sűrűségfüggvény értelmezéséből következik az alábbi két alapvető tulajdonsága:

•

$$f(x) \geq 0$$

•

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

8.1. Várható érték, szórás

8.2. Definíció.

Egy folytonos valószínűségi változó várható értéke:

$$E[\xi] = \int_{-\infty}^{+\infty} x \cdot f(x)dx$$

8.1. Megjegyzés.

A parciális integrálás módszere szerint:

$$E[\xi] = \int_{-\infty}^{+\infty} x \cdot f(x)dx = \int_{-\infty}^{+\infty} (1 - F(x)) dx$$

8.3. Definíció.

Egy folytonos valószínűségi változó szórásnégyzete:

$$D^2[\xi] = \int_{-\infty}^{+\infty} (x - E[\xi])^2 \cdot f(x)dx = \int_{-\infty}^{+\infty} x^2 \cdot f(x)dx - E^2[\xi]$$

8.2. Nevezetes folytonos eloszlások

(i) Az $[a, b]$ intervallumon **egyenletes eloszlású** valószínűségi változó sűrűségfüggvénye:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{ha } a \leq x \leq b \\ 0 & \text{egyébként} \end{cases} .$$

Az eloszlásfüggvény:

$$F(x) = \begin{cases} 0 & \text{ha } x < a \\ \frac{x-a}{b-a} & \text{ha } a \leq x \leq b \\ 1 & \text{ha } b < x \end{cases}$$

Várható értéke és szórása:

$$E = \frac{a+b}{2} \quad D = \frac{b-a}{\sqrt{12}}.$$

ábra

(ii) Az **exponenciális eloszlású** valószínűségi változó sűrűségfüggvénye:

$$f(x) = \begin{cases} 0 & \text{ha } x < 0 \\ \lambda e^{-\lambda x} & \text{ha } x \geq 0 \end{cases}$$

ahol $\lambda > 0$ paraméter.

Az eloszlásfüggvény:

$$f(x) = \begin{cases} 0 & \text{ha } x < 0 \\ 1 - e^{-\lambda x} & \text{ha } x \geq 0 \end{cases}$$

Várható értéke és szórása:

$$E = \frac{1}{\lambda} \quad D = \frac{1}{\lambda}.$$

ábra

(iii) A **gamma eloszlású** valószínűségi változó sűrűségfüggvénye:

$$f(x) = \begin{cases} 0 & \text{ha } x < 0 \\ \frac{\lambda^p}{\Gamma(p)} x^{p-1} e^{-\lambda x} & \text{ha } x \geq 0 \end{cases}$$

ahol $\lambda > 0, p > 0$ paraméterek, valamint a gamma függvény:

$$\Gamma(p) = \int_0^{\infty} x^{p-1} e^{-x} dx.$$

A gamma eloszlású valószínűségi változó várható értéke és szórása:

$$E = \frac{p}{\lambda} \quad D = \frac{\sqrt{p}}{\lambda}.$$

ábra

8.2. Megjegyzés.

A gamma eloszlás speciális esete az exponenciális eloszlás. ($p = 1$ eset.)

8.3. Megjegyzés.

(iv) A χ^2 **eloszlású** valószínűségi változó is a gamma eloszlás speciális esete.
($p = \frac{n}{2}$, $\lambda = \frac{1}{2}$ eset.) Statisztikai alkalmazások miatt fontos. Sűrűségfüggvénye tehát:

$$f(x) = \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} \quad \text{ha} \quad x \geq 0$$

Várható értéke és szórása:

$$E = n \quad D = \sqrt{2n}.$$

ábra

(v) A **normális eloszlású** vagy más néven Gauss eloszlású valószínűségi változó sűrűségfüggvénye:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

ahol m tetszőleges, $\sigma > 0$ paraméterek.

Várható értéke és szórása:

$$E = m \quad D = \sigma.$$

ábra

Fontos az $m = 0$, $\sigma = 1$ speciális eset, amit *standard normális* eloszlásnak neveznek, ugyanis tetszőleges normális eloszlás eloszlásfüggvényének kiszámítása a standard normálisra való visszavezetéssel történik:

Ha ξ egy m várható értékű és σ szórású normális eloszlású valószínűségi változó, akkor

$$F(x) = P(\xi < x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(t-m)^2}{2\sigma^2}} dt$$

Az integrált az

$$u = \frac{t-m}{\sigma}$$

helyettesítéssel átalakítva kapjuk, hogy

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-m}{\sigma}} e^{-\frac{u^2}{2}} du = \Phi\left(\frac{x-m}{\sigma}\right).$$

Itt tehát Φ a standard normális eloszlásfüggvénye. Az eljárást szokás sztenderdizálásnak nevezni.

A normális eloszlás kiemelkedő jelentőségű a *centrális határeloszlás tétel* miatt.

8.1. Tétel. Centrális határeloszlás tétel

Ha $\xi_1, \xi_2, \dots, \xi_n$ független azonos eloszlású valószínűségi változók m várható értékkel és véges d szórással, akkor az összegük határeloszlása normális, mégpedig

$$\lim_{n \rightarrow \infty} P \left(\frac{\xi_1 + \xi_2 + \dots + \xi_n - nm}{d\sqrt{n}} < x \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du = \Phi(x)$$

Még két fontos tétel bizonyítás nélkül:

8.2. Tétel.

Véges sok független normális eloszlású valószínűségi változó összege is normális eloszlású. Az összeg várható értéke, a várható értékek összege. Az összeg szórásnégyzete a szórásnégyzetek összege.

8.3. Tétel.

Véges sok független standard normális eloszlású valószínűségi változó négyzetösszege χ^2 eloszlású valószínűségi változó.

9. Statisztika

A statisztika a matematikának azon része, amely módszereket szolgáltat adatok gyűjtéséhez, rendszerezéséhez, feldolgozásához. Feladata, hogy megfigyelésekből, mérésekből származó adathalmazból információt olvasson ki. Segítsen azok értelmezésében és bemutatásában.

9.1. Statisztikai minta

Tipikus esetben vagy egy nagyobb populációról vagy méréssel megfigyelt jelenségről kívánunk információt szerezni. Nagyméretű populáció esetén kiválasztjuk annak egy részét - ez a minta - és azon végezzük a megfigyeléseket, vagy méréseket. A megfigyelések vagy mérések eredménye a statisztikai minta.

A statisztikai minták két nagy csoportja: Kvantitatív, vagyis nem számokkal jellemzett mennyiségek halmaza.

Kvalitatív, vagyis numerikusan jellemzett mennyiségek. Mi a továbbiakban erre fordítjuk figyelmünket.

9.2. Mintavételi módszerek

A mintavételnek két alaptípusa van: **véletlen mintavétel** és **nem véletlen mintavétel**. Ez utóbbi során a megmintázott populáció nem minden elemének van esélye bekerülni a mintába. A minta nagymértékben függ a döntési helyzetben levő személytől, amely szubjektivitásra, torzításra ad lehetőséget és ezért matematikai módszerek nem alkalmazhatók.

A továbbiakban természetesen csak a véletlen mintavétellel foglalkozunk. Ennek három típusát ismertetjük.

Egyszerű véletlen minta.

Egyszerű véletlen statisztikai mintáról beszélünk, ha egy N elemű sokaságból (populáció) úgy választunk ki egy n elemű mintát, hogy a populáció bármely eleme ugyanakkora valószínűséggel kerülhet bele a mintába.

Rétegezett mintavétel

Rétegezett mintavétel esetén először a teljes sokaságot rétegekre bontjuk. A rétegeknek nem kell azonos méretűnek lenni, de nem fedhetik át egymást és le kell fedniük az egész populációt. Ezután minden rétegből egyszerű mintát veszünk. Ezzel a módszerrel elérhető, hogy a populáció különböző elemei különböző valószínűséggel kerüljenek bele a mintába.

Klaszterezett mintavétel

Klaszterezett mintavétel során a populációt klaszterekre bontjuk, majd a klaszterek közül véletlenszerűen választunk néhányat. A klasztereknek sem kell azonos méretűnek lenni, de a kiválasztott klaszter minden elemét feldolgozzuk.

A kiválasztott mintán azután a kérdéses feladattól függően *változókat* értelmezhetünk. Általában két eset szokott előfordulni: az egyik, amikor csak azt figyeljük, hogy a kiválasztott mintaelem rendelkezik-e a vizsgálat szempontjából érdekes tulajdonsággal, vagy sem. Ekkor azt szeretnénk tudni, hogy a teljes populáció hányad része rendelkezik ezzel a tulajdonsággal. Az erre irányuló becslést szokás **aránybecslésnek** nevezni.

Más esetben a mintaelemekhez (vagy a teljes populáció elemeihez) számszerűen jellemezhető változót rendelünk hozzá. Mivel ezen értékek megjelenése, mintába kerülése véletlentől függ, ezért ezt természetes módon valószínűségi változóval modellezzük. Ilyenkor ennek a valószínűségi változónak a tulajdonságaira vagyunk kíváncsiak. A valószínűségi változó megfigyelt (mért) értékei alkotják a **statisztikai mintát**.

Jelölje ξ a megfigyelni kívánt valószínűségi változót. Végezzünk n -szer egymástól függetlenül, azonos körülmények között megfigyeléseket. Legyen ezek eredménye x_1, x_2, \dots, x_n . Ez a statisztikai minta, amelyből következtetni kívánunk a ξ valószínűségi változóra.

9.3. Empirikus eloszlásfüggvény

Legyen x_1, x_2, \dots, x_n a statisztikai minta. Rendezzük ezen mintaelemeket nagyság szerint: $x_1^* < x_2^* < \dots < x_n^*$. Definiáljuk most az $F_n(x)$ lépcsősfüggvényt a következő módon:

$$F_n(x) = \begin{cases} 0 & \text{ha } x \leq x_1^* \\ \frac{k}{n} & \text{ha } x_k^* < x \leq x_{k+1}^* \\ 1 & \text{ha } x_n^* < x \end{cases}$$

Ez az empirikus eloszlásfüggvény. Azért nevezik lépcsősfüggvénynek, mert a rendezett mintaelemeknél $\frac{1}{n}$ ugrása van.

ábra

Jelölje $F(x)$ a ξ valószínűségi változó eloszlásfüggvényét, azaz $F(x) = P(\xi < x)$. Az alábbi Glivenkotól származó tételt nagy jelentősége miatt a statisztika alaptételének is nevezik:

9.1. Tétel.

Legyen D_n az empirikus $F_n(x)$ eloszlásfüggvény és az $F(x)$ elméleti eloszlásfüggvény közötti legnagyobb eltérés:

$$D_n = \sup_{-\infty < x < +\infty} |F_n(x) - F(x)|.$$

Ez az eltérés 1 valószínűséggel zérushoz tart, ha n tart végtelenhez:

$$P(\lim_{n \rightarrow \infty} D_n = 0) = 1.$$

Kolmogorov vizsgálta azt a kérdést, hogy D_n milyen gyorsan tart zérushoz. Meghatározta $\sqrt{n}D_n$ eloszlását. Tőle származik a következő tétel:

9.2. Tétel.

$$\lim_{n \rightarrow \infty} P(\sqrt{n}D_n < z) = \sum_{i=-\infty}^{+\infty} (-1)^i e^{-2i^2 z^2} = K(z).$$

9.1. Megjegyzés.

A $K(z)$ függvény értékeit nem könnyű meghatározni, ezért vagy már előre elkészített táblázatból vesszük az értékeit, vagy számítógépet használunk.

Mivel $K(z)$ határeloszlás-függvény, ezért nagy mintaelemszám esetén alkalmazható. Kis mintákra módosított értékekből készítenek táblázatot.

Az elméleti és empirikus eloszlásfüggvények közötti egyoldalú eltéréseket Szmirnov vizsgálta. Tőle származik a következő tétel:

9.3. Tétel.

Ha

$$D_n^+ = \sup_{-\infty < x < +\infty} (F_n(x) - F(x))$$

$$D_n^- = \sup_{-\infty < x < +\infty} (F(x) - F_n(x))$$

akkor

$$\lim_{n \rightarrow \infty} P(\sqrt{n}D_n^+ < z) = \lim_{n \rightarrow \infty} P(\sqrt{n}D_n^- < z) = 1 - e^{-2z^2}.$$

10. Paraméterbecslés

10.1. Pontbecslés

A statisztika alaptétele, valamint Kolmogorov és Szmirnov tételei mutatják, hogy az empirikus eloszlásfüggvény közel van az elméleti eloszlásfüggvényhez. Következés képpen az empirikus eloszlásból számolt paramétereknek is közel lenni a modellben szereplő ξ valószínűségi változó paramétereire.

Az empirikus eloszlás várható értéke:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Az empirikus eloszlás szórásnégyzete:

$$s_n^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \left(\sum_{i=1}^n x_i^2 \right) - \bar{x}^2.$$

Általában, ha $\theta(\xi)$ a ξ valószínűségi változó valamely paramétere, akkor remélhetőleg a statisztikai mintából számított $\hat{\theta}_n = \hat{\theta}(x_1, x_2, \dots, x_n)$ a $\theta(\xi)$ becslése.

10.1. Definíció.

Azt mondjuk, hogy a $\hat{\theta}(x_1, x_2, \dots, x_n)$ a $\theta(\xi)$ paraméter **torzítatlan** becslése, ha $\hat{\theta}_n$ várható értéke éppen θ , azaz

$$E \left[\hat{\theta}(x_1, x_2, \dots, x_n) \right] = \theta(\xi).$$

10.1. Tétel.

Az empirikus eloszlás várható értéke torzítatlan becslése az elméleti várható értéknek:

$$E[\bar{x}] = E[\xi]$$

Az empirikus eloszlás szórásnégyzete nem torzítatlan becslése az elméleti szórásnégyzetnek. Igaz ugyanis a következő állítás:

10.2. Tétel.

$$E[s_n^2] = \frac{n-1}{n} D^2[\xi]$$

Következés képpen az

$$s_n^{*2} = \frac{n}{n-1} s_n^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

torzítatlan becslése a szórásnégyzetnek és így

$$s_n^* = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$

úgynevezett **korrigált szórás** torzítatlan becslése az elméleti szórásnak.

10.1. Megjegyzés.

Ha $\hat{\theta}_n$ torzítatlan becslése a θ paraméternek, vagyis $E[\hat{\theta}_n] = \theta$, akkor $E[\hat{\theta}_n - \theta] = 0$.

10.2. Definíció.

Ha $\hat{\theta}_n$ torzítatlan becslése a θ paraméternek és még az is teljesül, hogy

$$\lim_{n \rightarrow \infty} E[(\hat{\theta}_n - \theta)^2] = 0,$$

akkor azt mondjuk, hogy $\hat{\theta}_n$ **konzisztens becslése** a θ paraméternek.

10.3. Tétel.

Az \bar{x} konzisztens becslése a $E[\xi]$ várható értéknek, mert

$$D^2[\bar{x}] = \frac{1}{n} D^2[\xi].$$

10.2. Konfidencia intervallum

Az a kijelentés, hogy a statisztikai mintából számított $\hat{\theta}_n$ torzítatlan becslése a θ paraméternek nem elég információ. Azt is szeretnénk tudni, hogy mennyire jó becslése a θ paraméternek. Rendszerint úgy járunk el, hogy egy intervallumot adunk meg, amelybe nagy valószínűséggel benne van a becsülendő paraméter igazi értéke. Ennek neve **konfidencia intervallum**, vagy megbízhatósági intervallum. Természetesen a konfidencia intervallum hossza függ a választott valószínűség értéktől, azaz a becslés **megbízhatósági szintjétől**. Nagyobb megbízhatósági szinthez szélesebb konfidencia intervallum tartozik. A műszaki gyakorlatban leggyakrabban a 95%-os vagy 99%-os megbízhatósági szintet használják. A 95%-os megbízhatósági szint azt jelenti, hogy a tévedés valószínűsége, vagyis hogy a paraméter tényleges értéke a konfidencia intervallumon kívül esik, legfeljebb 0,05.

A becslések minden esetben arra épülnek, hogy az egyszerű mintavétel feltételei teljesülnek, azaz a mintaelemek egymástól függetlenek, azonos eloszlásúak és azonos valószínűséggel vannak kiválasztva. Ezért igaz rájuk a (8.1) centrális határeloszlás tétel, illetve az alábbi tétel:

10.4. Tétel. Nagy számok törvényének Bernoulli féle alakja

Ha valamely p valószínűségű esemény n számú független kísérlet során k -szor következik be, akkor a $\frac{k}{n}$ relatív gyakoriság jól közelíti a p valószínűséget olyan értelemben, hogy tetszőleges $\epsilon > 0$ szám esetén

$$P\left(\left|\frac{k}{n} - p\right| \geq \epsilon\right) \leq \frac{pq}{n\epsilon^2} \leq \frac{1}{4n\epsilon^2},$$

vagy ugyan ez az ellentett eseményre fogalmazva:

$$P\left(\left|\frac{k}{n} - p\right| < \epsilon\right) > 1 - \frac{pq}{n\epsilon^2} \geq 1 - \frac{1}{4n\epsilon^2}.$$

Lássuk most részletesen a 9.2 pontban említett mintavételi eljárásokat.

Egyszerű mintavétel, átlag becslése

Tegyük fel, hogy van egy N elemű populáció, ahol az elemekhez valamilyen módon számok vannak hozzárendelve. Ezen számok átlagát szeretnénk becsülni. (Pl. a populáció a BME hallgatói és azt szeretnénk tudni, hogy havonta mennyi pénzt költenek sportolásra. Mindenkit kikérdezni hosszadalmas volna, ezért egy mintából akarjuk becsülni az átlagot.) Válasszunk ki véletlenszerűen egy n elemű mintát. Ekkor az átlag becslött értéke:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

A konfidencia intervallum meghatározásához szükség van a mintaelemek korrigált szórására:

$$s_n^* = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$

Ennek segítségével a konfidencia intervallum:

$$\bar{x} \pm t \frac{s_n^*}{\sqrt{n}} \sqrt{\frac{N-n}{N}}$$

ahol a t a Student-féle t -eloszlás megfelelő értéke. Pl. 95 %-os megbízhatósági szint esetén, ha a minta nagysága kb. 20 és 50 közé esik akkor $t \approx 2$, ha pedig $n > 50$, akkor $t = 1,96$.

10.2. Megjegyzés.

Ha N sokkal nagyobb, mint n , akkor a $\sqrt{\frac{N-n}{N}} \approx 1$, ezért ez a korrekciós faktor elhagyható.

10.1. Példa.

Bizonyos fajta térburkoló követ raklapokon árulnak. Az építésvezető az első 10 raklap felhasználásáról a következőket jegyezte fel:

4,79 m ²	5,08 m ²	5,17 m ²	4,83 m ²	4,95 m ²
5,13 m ²	5,05 m ²	4,88 m ²	5,07 m ²	5,06 m ²

Adjunk becslést arra, hogy átlagosan hány m² lefedésére alkalmas egy raklap!

Adjunk 95%-os konfidencia intervallumot az átlagos területre! Ha az építésvezetőnek 500m² területet kell burkoltatni, hány raklapot rendeljen?

Megoldás:

A táblázatban levő adatok számtani közepe: $\bar{x} = 5,001m^2$.

A táblázatban levő adatok szórása: $s_n^* = 0,130422m^2$.

Ilyen kis minta esetén a t értékét a Student-féle eloszlás táblázatából vesszük, ami jelen esetben $t = 2,228$.

Így a konfidencia intervallum $5,001 \pm 0,087174$, azaz $(4,9138; 5,0881)$.

Mivel egy raklap kb. $5m^2$ lefedésére elég, ezért $500m^2$ lefedéséhez kb. 100 raklap kell.

Az építésvezető nem akarja megkockáztatni, hogy a munka végén egy kevés burkolóanyag hiányozzon - úgy gondolja, hogy kisebb veszteség, ha marad egy kevés burkolóanyag, mint ha nem végeznek határidőre, ezért többet rendel. Elegendő-e mondjuk 102 raklapot rendelni?

Erre először általánosságban keressük a választ:

Ha egy N elemű populációból választottunk n elemű mintát és abban a várható érték

$$\bar{x} = \frac{\sum x_i}{n} \quad \text{és a szórás} \quad s_n^* = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

akkor ezt az egész populációra kivetítve azt kapjuk, hogy

$$\hat{x} = N\bar{x}$$

amelyhez a konfidencia intervallum:

$$\hat{x} \pm t \cdot N \cdot \frac{s_n^*}{\sqrt{n}} \sqrt{\frac{N - n}{N}}$$

Visszatérve a példánkra, ha

$$N = 102, \quad n = 10, \quad t = 1,96 \quad \bar{x} = 5,001m^2 \quad \text{és} \quad s_n^* = 0,130422m^2$$

akkor a konfidencia intervallum (95%-os szinten) $510,102 \pm 7,8694$, azaz $(502,23; 517,97)$. Tehát nagy valószínűséggel elég a 102 raklap.

Egyszerű mintavétel, aránybecslés

Most is tegyük fel, hogy van egy N elemű populáció, amelyben bizonyos elemek rendelkeznek valamely tulajdonsággal. Ezen elemek arányát szeretnénk becsülni. (Pl. a populáció bizonyos építőanyag fajta. Azt szeretnénk tudni, mennyi selejtes van benne, de az egész tételt átvizsgálni nem gazdaságos. Ezért mintát veszünk belőle és abból fogjuk becsülni a selejtarányt.)

Válasszunk ki véletlenszerűen egy n elemű mintát. Ha a mintában k elem rendelkezik a kérdéses tulajdonsággal, akkor a populációbeli arányra a becslés:

$$\bar{p} = \frac{k}{n}$$

A megfelelő konfidencia intervallum:

$$\bar{p} \pm t \sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \sqrt{\frac{N-n}{N}}$$

Ha most arra is szeretnénk becslést, hogy a populációban hányan rendelkeznek a kérdéses tulajdonsággal, akkor a mintából kapott becslést kivetítjük a populációra, így a becslés:

$$\hat{p} = N \cdot \bar{p}.$$

Ehhez a konfidencia intervallum:

$$N \cdot \bar{p} \pm t \cdot N \sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \sqrt{\frac{N-n}{N}}$$

10.2. Példa.

Tegyük fel, hogy az építésvezető rendelt 5000 db. másodosztályú csempét. 100 db. átvizsgálásakor 4 db. hibásat találtak. Szeretné tudni, hogy várhatóan hány hibás van az egész tételben és 95% valószínűséggel legfeljebb hány darabot kell kidobni.

Megoldás: Most az adatok:

$$N = 5000, \quad n = 100, \quad t = 1,96 \quad \text{és} \quad k = 4$$

ezért $p = 0,04$ a szórás=0,01959 a korrekciós faktor=0,9899 és így a selejtarány konfidencia intervallumának fele=0,03802. Kivetítve az egész populációra kapjuk, hogy a selejtek várható száma = 200 az erre vonatkozó konfidencia intervallum fele ≈ 190 a konfidencia intervallum: (10; 390). Vagyis legfeljebb 5% a valószínűsége, hogy több mint 390 hibás darabot találunk.

Rétegezett mintavétel

Jelölések:

- r = a rétegek száma
- N_i = az elemek száma az i -dik rétegben
- N = $N_1 + N_2 + \dots + N_r$ az elemek száma a teljes populációban
- n_i = az i -dik rétegből kiválasztott mintaelemek száma
- n = $n_1 + n_2 + \dots + n_r$ az összes mintaelemek száma
- \bar{x}_i = az i -dik rétegből kiválasztott mintaelemek átlaga
- s_i = az i -dik rétegből kiválasztott mintaelemek szórása
- p_i = az i -dik rétegből kiválasztott mintaelemek számított arány,

azaz ha n_i mintaelem közül k_i darab rendelkezik a megfigyelt tulajdonsággal, akkor $p_i = \frac{k_i}{n_i}$.

A várható érték becslése:

$$\bar{x} = \frac{1}{N} (N_1\bar{x}_1 + N_2\bar{x}_2 + \dots + N_r\bar{x}_r)$$

Az ehhez tartozó konfidencia intervallum:

$$\bar{x} \pm t \sqrt{\frac{1}{N^2} \sum N_i^2 \frac{s_i^2}{n_i} \left(\frac{N_i - n_i}{N_i} \right)}$$

A paraméter értékét a teljes populációra vetítve a

$$\bar{\tau} = N\bar{x} = N_1\bar{x}_1 + N_2\bar{x}_2 + \dots + N_r\bar{x}_r$$

becslés adódik, amelynek konfidencia intervalluma

$$\bar{\tau} \pm t \sqrt{\sum N_i^2 \frac{s_i^2}{n_i} \left(\frac{N_i - n_i}{N_i} \right)}$$

Aránybecslésre a

$$\bar{p} = \frac{1}{N} (N_1p_1 + N_2p_2 + \dots + N_rp_r)$$

kifejezést használjuk, amelynek konfidencia intervalluma:

$$\bar{p} \pm t \sqrt{\frac{1}{N^2} \sum N_i^2 \frac{p_i(1-p_i)}{n_i-1} \left(\frac{N_i - n_i}{N_i} \right)}.$$

Klaszterezett mintavétel

Jelölések:

- N = a klaszterek száma
- n = a mintába beválasztott klaszterek száma
- m_i = az i -ik klaszterben az elemek száma
- M = a populáció elemszáma, azaz $M = \sum_{i=1}^N m_i$
- \bar{M} = $\frac{M}{N}$ = a klaszterek átlagos mérete a populációban
- \bar{m} = a klaszterek átlagos mérete a mintában
- x_i = az i -ik klaszterbeli megfigyelések összege

A várható érték becslése:

$$\bar{x} = \frac{\sum x_i}{\sum m_i}$$

és a konfidencia intervallum:

$$\bar{x} \pm t \sqrt{\left(\frac{N-n}{Nn\bar{M}^2} \right) \sum \frac{(x_i - \bar{x}m_i)^2}{n-1}}$$

A teljes érték becslése:

$$\bar{\tau} = M\bar{x} = M \frac{\sum x_i}{\sum m_i}.$$

A konfidencia intervallum:

$$\bar{\tau} \pm t \sqrt{N^2 \left(\frac{N-n}{Nn} \right) \sum \frac{(x_i - \bar{x}m_i)^2}{n-1}}.$$

Az arány becslése:

$$\bar{p} = \frac{\sum k_i}{\sum m_i}$$

ahol k_i az i -ik klaszterben azon elemek száma, amelyek a megfigyelt tulajdonsággal rendelkeznek. Az ehhez tartozó konfidencia intervallum:

$$\bar{p} \pm t \sqrt{\left(\frac{N-n}{Nn\bar{M}^2} \right) \sum \frac{(k_i - \bar{p}m_i)^2}{n-1}}.$$

Mekkora legyen a minta?

Jelölje d a kívánt konfidencia intervallum felének a hosszát, valamint legyen $w_i = \frac{n_i}{n}$ az i -ik rétegben levő mintaelemek aránya az egész mintához.

A szükséges mintaelemszám

Egyszerű véletlen minta	
Átlag becslése	$n = \frac{N s^2}{(N-1) \frac{d^2}{t^2} + s^2}$
Egész populációra	$n = \frac{N s^2}{(N-1) \frac{d^2}{t^2 N^2} + s^2}$
Arány	$n = \frac{N p q}{(N-1) \frac{d^2}{t^2} + p q}$
Rétegzett minta	
Átlag becslése	$n = \frac{\sum \left(\frac{N_i^2 s_i^2}{w_i} \right)}{N^2 \frac{d^2}{t^2} + \sum N_i s_i^2}$
Egész populációra	$n = \frac{\sum \left(\frac{N_i^2 s_i^2}{w_i} \right)}{\frac{d^2}{t^2} + \sum N_i s_i^2}$
Arány	$n = \frac{\sum \left(\frac{N_i^2 p_i q_i}{w_i} \right)}{N^2 \frac{d^2}{t^2} + \sum N_i p_i q_i}$
Klaszterezett mintavétel	
Átlag becslése	$n = \frac{N s_c^2}{N \frac{\bar{M}^2 d^2}{t^2} + s_c^2}$
Egész populációra	$n = \frac{N s_c^2}{N \frac{d^2}{t^2 N^2} + s_c^2}$
Arány	$n = \frac{N s_c^2}{N \frac{\bar{M}^2 d^2}{t^2} + s_c^2}$

11. Hipotézisvizsgálat

A *hipotézis* egy, a populációra vonatkozó állítás. Két tipikus eset van; az egyik, amikor a hipotézis arra vonatkozik, hogy a populáció paramétere bizonyos értéket (vagy értékeket) vesz fel. A másik eset, amikor a populációval kapcsolatos valószínűségi változók eloszlására vonatkozik az állítás. A hipotézisvizsgálat során a populációból vett minta alapján döntünk arról, hogy elfogadható-e az állítás, vagy elvetendő. Az eljárást *statisztikai próbának* nevezik. Mi itt csak az első esettel, a paraméteres próbával foglalkozunk.

11.1. Példa.

Egy szál aszbeszigetelő anyagot gyártó cég új termékkel jelentkezik a piacon. Azt állítja, hogy a termékének a hővezetési tényezője $k = 0,03666 [W/mK]$. A fogyasztóvédelmi hatóság 81

elemű mintán elvégezve a méréseket ellenőrizni akarja a gyártó állítását. Az ellenőrző mérések szerint a mintadarabok átlagos hővezetési tényezője $k = 0,03670[W/mK]$ és szórása $s = 0,002$. Mi legyen az álláspontjuk?

Megoldás:

A hipotézis vizsgálat 5 lépésből áll:

1) Először megfogalmazzuk a hipotézist és az ellenhipotézist. A (11.1) példa szerint a *null-hipotézis*, azaz amit vizsgálni szeretnénk:

$$H_0 : \text{a populáció hővezetési tényezőjének várható értéke } \leq 0,03666[W/mK]$$

Az ellenhipotézis:

$$H_1 : \text{a populáció hővezetési tényezőjének várható értéke } > 0,03666[W/mK]$$

Megjegyzés: A feladattól függően lehet egyoldalú ellenhipotézist vizsgálni, mint az itteni példában, vagy lehet kétoldalú $H_1 \neq 0,03666$ ellenhipotézist, ha arra van szükség.

2) A szignifikancia szint megválasztása. Mivel a döntés egy véletlentől függő statisztikai mintára épül, ezért ki van téve a véletlen ingadozásnak. Arra törekszünk, hogy a hibás döntés valószínűsége kicsi legyen. Választunk egy kis ϵ számot (gyakorlatban sokszor $\epsilon = 0,05$, vagy $\epsilon = 0,01$) és úgy konstruáljuk meg a döntést, hogy ha igaz a H_0 null-hipotézis, akkor nagy $1 - \epsilon$ valószínűséggel el is fogadjuk igaznak. Valójában kétféle hiba történhet:

Elsőfajú hiba: Igaz a H_0 null-hipotézis, mégis elutasítjuk. Ennek valószínűsége legyen kisebb mint ϵ .

Másodfajú hiba: Nem igaz a H_0 null-hipotézis, mégis elfogadjuk.

	elfogadjuk a H_0 hipotézist	elutasítjuk a H_0 hipotézist
H_0 igaz	helyes döntés	elsőfajú hiba
H_0 hamis	másodfajú hiba	helyes döntés

3) Megválasztjuk a próbastatisztikát. A (11.1) példa esetén értelemszerűen a várható érték becslése a mintaátlag alapján, a szórás becslése a minta korrigált szórása alapján történhet. (Lásd t-próba lejjebb.)

Egyszerűbb a helyzet, ha a szórás ismert. Ekkor a centrális határeloszlás tételből (8.1) tudjuk, hogy

$$z = \sqrt{n} \cdot \frac{E[X] - \bar{x}}{\sigma} \tag{83}$$

standard normális eloszlású.

4) A döntési szabály megalkotása. Ez az elfogadási, illetve elutasítási tartományok kijelölését jelenti. Ha a 3) pontban megválasztott próbastatisztika értéke az elfogadási tartományba esik, akkor elfogadjuk a H_0 hipotézist, ha az elutasítási tartományba esik, akkor elutasítjuk.

Azt a pontot, amelyik elválasztja az elfogadási és elutasítási tartományt *kritikus értéknek* nevezük.

Nézzük most a (11.1) példán keresztül, hogy ez hogyan történhet.

$$H_0 : \text{a populáció hővezetési tényezőjének várható értéke} \leq 0,03666[W/mK]$$

Az ellenhipotézis:

$$H_1 : \text{a populáció hővezetési tényezőjének várható értéke} > 0,03666[W/mK]$$

Ha a minta átlaga kisebb volna mint $0,0366[W/mK]$, akkor persze semmi okunk nem volna a H_0 hipotézis elutasítására, hiszen annak megfelelő esemény következett be. Példánkban azonban $\bar{x} = 0,037 > 0,0366$. Most az a kérdés, hogy olyan nagy-e a mintából számolt átlag, hogy az valószínűtlen (kisebb a valószínűsége mint $\epsilon = 0,05$), ha igaz a H_0 hipotézis. Nagy minta esetén - 81 elemű minta nagynak számít - az elméleti szórás helyettesíthető a tapasztalati szórással, tehát használhatjuk a (83) összefüggést. Ha már megválasztottuk az $\epsilon = 0,05$ értéket, akkor a standard normális eloszlásfüggvény értékeiből tudjuk, hogy $\Phi(1,65) \approx 0,95$. Ezért az 5%-os szignifikancia szinten (95%-os megbízhatósági szinten) a statisztika értékének kisebbnek kell lenni, mint $\bar{x} + 1,65 \cdot \frac{\sigma}{\sqrt{n}} = 0,0366 + 1,65 \cdot \frac{0,002}{9} = 0,03696$. Ez a $0,03696$ a *kritikus érték*. Ha a mért eredmény ennél kisebb, nincs okunk elutasítani a H_0 hipotézist, ha nagyobb, akkor elutasítjuk. Jelen esetben nagyobb, tehát elutasítjuk a H_0 hipotézist.

ábra

Ha a (11.1) példában azt szeretnénk ellenőrizni statisztikai vizsgálattal, hogy a

$$H_0 : \text{a populáció hővezetési tényezőjének várható értéke} = 0,03666[W/mK]$$

hipotézis elfogadható-e, a

$$H_1 : \text{a populáció hővezetési tényezőjének várható értéke} \neq 0,03666[W/mK]$$

ellenhipotézissel szemben, akkor a döntési tartományt az elméleti értékre szimmetrikusan választjuk meg, azaz ha \bar{x} az $E[X] \pm z \cdot \frac{\sigma}{\sqrt{n}}$ intervallumba esik elfogadjuk a H_0 hipotézist, ha pedig kívül esik, elutasítjuk. Ebben az esetben 95%-os megbízhatósági szinthez a $z = 1,96$ érték tartozik. Ez a két oldalú próba, szemben az előbbi egy oldalú próbával. Így a példabeli elfogadási tartomány

$$0,0366 \pm 0,00043 = (0,03616; 0,03703)$$

intervallum, amelybe a statisztika értéke bele esik, tehát nincs elegendő ok a null-hipotézis elutasítására.

ábra

5) A döntés. Utolsó lépésként a mintából számított statisztika, a szignifikancia szint és a már kijelölt elfogadási tartomány alapján meghozzuk a döntést.

11.1. Megjegyzés.

Sok statisztikai programcsomag úgy működik, hogy meg kell adni a következő információkat:

- A H_0 hipotézis
- Az ellenhipotézis egy- vagy kétoldalú
- A használni kívánt statisztika
- A statisztikai minta, amelyből a program kiszámítja a statisztika aktuális értékét.

A program egy p -értékkel tér vissza. Ez annak valószínűsége, hogy ha igaz a H_0 hipotézis, akkor milyen valószínűséggel következhet be a statisztika értéke. Ha a p -érték kicsi, kisebb mint az ϵ szignifikancia szint, akkor nagyon kis valószínűségű esemény következett be, tehát elvetjük a null-hipotézist.

Például a (83) példa esetén az egyoldalú ellenhipotézissel a program kiszámítaná a

$$z = \sqrt{n} \cdot \frac{E[X] - \bar{x}}{\sigma} = \sqrt{81} \cdot \frac{0,037 - 0,0366}{0,002} = 1,8$$

értékhez tartozó $p = 1 - \Phi(1,8)$ valószínűséget, amely 0,0359. Ez kisebb, mint 0,05 ezért a null-hipotézist elutasítjuk. Kétoldalú ellenhipotézis megadása esetén a program a

$$p = 1 - (\Phi(1,8) - \Phi(-1,8)) = 2 - 2\Phi(1,8) = 0,0718$$

értéket adja, amely nagyobb, mint a választott szignifikancia szint, tehát nem utasítjuk el a H_0 hipotézist.

11.1. u-próba, t-próba

11.1.1. Egymintás próba várható értékre

Tegyük fel, hogy a populáción vizsgált ξ valószínűségi változó normális eloszlású. Legyen a H_0 hipotézis az, hogy a várható értéke $E[\xi] = \mu$ és n a minta nagysága. Tekintsük a

$$\sqrt{n} \cdot \frac{\bar{x} - \mu}{s} \quad (84)$$

statisztikát.

- Ha az s szórás ismert, vagy a minta nagysága elég nagy (praktikusan $n > 30$), akkor a (84) statisztika normális eloszlású.
- Ha az s szórás nem ismert és ezért a mintából becsüljük, továbbá a minta nem elég nagy (praktikusan $n < 30$), akkor a (84) statisztika *Student-féle t-eloszlású* $n - 1$ szabadsági fokkal.

11.2. Példa.

11.3. Példa.

11.1.2. Kétmintás próba várható értékre

Tegyük fel, hogy a populáción vizsgált ξ és η valószínűségi változók függetlenek és normális eloszlásúak. Legyen a H_0 hipotézis az, hogy a várható értéke $E[\xi] = E[\eta]$. Tekintsük a

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (85)$$

statisztikát, ahol

- \bar{x}_1 és \bar{x}_2 a mintaközépek
- n_1 és n_2 a minták nagysága
- s_1 és s_2 a minták szórása.

A (85) statisztika standard normális eloszlású, ha n_1 és n_2 elég nagy.

Ha n_1 és n_2 nem elég nagy, akkor

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad (86)$$

statisztika, ahol

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Student-féle t eloszlású $n_1 + n_2 - 2$ szabadsági fokkal.

11.4. Példa.

11.1.3. Párosított t-próba

Most a populáción vizsgált ξ és η valószínűségi változók nem függetlenek, de normális eloszlásúak. Azt a null-hipotézist kívánjuk tesztelni, hogy a két valószínűségi változó különbségének várható értéke zérus-e. (Lsd. pl. (11.5) példa). Ekkor a

$$t = \sqrt{n} \cdot \frac{\bar{d}}{s_d} \quad (87)$$

statisztika Student-féle t eloszlású $n - 1$ szabadsági fokkal, ahol

$$s_d = \sqrt{\frac{\sum d^2 - \frac{(\sum d)^2}{n}}{n - 1}}$$

Amennyiben n elég nagy, akkor itt is a t eloszlás helyett a normális eloszlás használható.

11.5. Példa.

11.1.4. Arány tesztelése

Legyen valamely tulajdonság aránya a teljes populáción π és az n elemű mintában p . Legyen a H_0 hipotézis az, hogy $\pi = p$. Ekkor a

$$t = \sqrt{n} \cdot \frac{p - \pi}{\sqrt{p(1-p)}}$$

statisztika $n - 1$ szabadsági fokú t eloszlású.

11.6. Példa.

11.1.5. Arányok egyezése

Azt vizsgáljuk, hogy két populációból vett független mintákban a kérdéses arány megegyezik-e.

- p_1 és p_2 az egyes populációkból vett mintában az arány
- n_1 és n_2 a minták nagysága
- k_1 és k_2 az egyes mintákban a kérdéses tulajdonsággal rendelkezők száma
- p_c közös arány: $p_c = \frac{k_1 + k_2}{n_1 + n_2}$.

Ekkor a

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}}$$

statisztika standard normális eloszlású.

11.7. Példa.

12. Illeszkedésvizsgálat

A műszaki gyakorlatban sokszor vagyunk olyan helyzetben, hogy nem ismerjük a szóban forgó valószínűségi változó eloszlását. Az előző pontban a paraméteres próbák ismertetése során feltettük, hogy a kérdéses valószínűségi változó normális eloszlású. Most azt szeretnénk ellenőrizni, hogy ez a feltevés elfogadható-e.

Először egy egyszerűbb esetet nézzünk.

12.1. χ^2 próba

Tegyük fel, hogy ξ egy diszkrét valószínűségi változó x_1, x_2, \dots, x_n lehetséges értékekkel és a hozzá tartozó valószínűségeloszlás: p_1, p_2, \dots, p_n , azaz $P(\xi = x_i) = p_i$. Végezzünk N független kísérletet. Azt tapasztaljuk, hogy a $\{\xi = x_1\}$ esemény ν_1 -szer, a $\{\xi = x_2\}$ esemény ν_2 -szer, ... a $\{\xi = x_n\}$ esemény ν_n -szer következett be. ($\sum_{i=1}^n \nu_i = N$.) Vizsgáljuk a $H_0 : P(\xi = x_i) = p_i$ hipotézist. Ha a H_0 hipotézis igaz, akkor a

$$\chi^2 = \sum_{i=1}^n \frac{(\nu_i - Np_i)^2}{Np_i} \quad (88)$$

statisztika $n - 1$ -ed fokú χ^2 eloszlású valószínűségi változó.

12.1. Példa.

folytonos eset

12.2. Példa.

12.2. Kolmogorov próba

szöveg

13. Függőség jellemzése

szöveg

13.1. Korreláció

szöveg

13.2. Regresszió

szöveg

Irodalomjegyzék

- [1] Ponomarjov, K.K. (1969) *Differenciálegyenletek felállítása és megoldása* Tankönyvkiadó, Budapest
- [2] Reimann József (1992) *Valószínűségelmélet és matematikai statisztika mérnököknek* Tankönyvkiadó, Budapest
- [3] Rényi Alfréd (1968) *Valószínűségszámítás* Tankönyvkiadó, Budapest
- [4] Rózsa Pál (1976) *Lineáris algebra és alkalmazásai* Műszaki Könyvkiadó, Budapest
- [5] Scharnitzky Viktor (1975) *Differenciálegyenletek* Műszaki Könyvkiadó, Budapest
- [6] Székely J. Gábor, Zsigri Gábor (1993) *Differenciálegyenletek és differenciálgeometria* BME jegyzeterv
- [7] Young, D.M. (1979) *Nagy lineáris rendszerek iterációs megoldása* Műszaki Könyvkiadó, Budapest