Nemlineáris egyenletrendszerek megoldása

Filep Illés Attila

2023. április 15.

Kivonat

Tartalomjegyzék

Bevezetés

A dokumentum célja a sztochasztikus folyamatok alkalmazása nevű tárgyon tanult, kiemelt elemek demonstrációja. A demonstráció MATLAB könyvtár elkészítésével történik. A könyvtárnak a célja, hogy szimbolikus matematikai eszközökkel a folyamatokat bemutassa. A könyvtárnak nem célja a semmilyen informatikai optimalizáltságot megvalósítani.

I. rész
Stacionárius folyamatok

1. Alapvető definíciók

1.1. Valószínűségi változó

Legyen:

- Ω egy nem üres halmaz
- $\{\omega : X(\omega) < x\} \in \mathcal{A}$
- $x \in \mathbb{R}$
- $\mathcal{A}$ az Ω részhalmazaiból alkotott esemény σ-algebrája (tehát $(\Omega, \mathcal{A})$ mérhető tér)

Akkor $X : \Omega \to \mathbb{R}$ függvényt valószínűségi változónak hívunk.

1.2. Sztochasztikus folyamat

A sztochasztikus folyamat (vagy véletlen folyamat) egy olyan matematikai modell, amely egy vagy több időfüggő véletlen változó által létrehozott folyamatot ír le. A sztochasztikus folyamatok olyan rendszerek leírására szolgálnak, amelyekben a jövő állapota részben véletlenszerűen határozza meg a múlt és a jelen állapotát.

A sztochasztikus folyamatok általában valószínűségi változók sorozataként jelennek meg, amelyeknek az idő függvényében változó értékei vannak. A folyamatot gyakran matematikailag leírt egyenletekkel vagy valószínűségi eloszlásokkal írják le.

A sztochasztikus folyamatok számos területen alkalmazhatók, például az anyag- és energiaátvitel, a kommunikációs rendszerek, a pénzügyek, az idősorok elemzése és a gépi tanulás területén.

1.2.1. Sztochasztikus folyamatok kompatibilitási feltételei

A sztochasztikus folyamatok kompatibilitási feltételei a következők:

- Az időpillanatok száma felsorolható, véges vagy végtelen, de számontartható.

- Az időpillanatok sorozata szigorúan növekvő, azaz $t_1 < t_2 < \cdots < t_n$ vagy $t_1 < t_2 < \cdots < t_\infty$.
- Az időpillanatok közötti időközök meghatározottak és végesek vagy végtelenek.
- A folyamat értékei véletlenszerűek, és általában valószínűségi változóként vannak definiálva.
- A folyamat értékei időfüggők, és az időbeli elmozdulásokkal szembeni szimmetriára vonatkozó korlátozásokat kell teljesítenie. Például a stacionárius folyamatok esetében az eloszlások nem változnak az idő múlásával, és az átlag és szórás időfüggetlen.

Ezen kívül a sztochasztikus folyamatoknál általában szükséges az ergodicitás feltétele, amely azt jelenti, hogy a folyamat minden pillanatban eléri minden lehetséges állapotát az idő végtelen futamán. Ez fontos feltétele a statisztikai tulajdonságok meghatározásának, mert lehetővé teszi a folyamat várható értékének becslését a mintavételezés révén.

1.3. Várható érték

Lényegében az első (centrális) momentum, egy funkcionál.

Diszkrét esetben

$$E(X) = \sum_{i=1}^{\infty} p_i x_i$$

Folytonos esetben

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

1.3.1. Várható érték létezésének a feltétele

Diszkrét esetben

$$E(X) = \sum_{i=1}^{\infty} p_i |x_i| < \infty$$

Folytonos esetben

$$E(X) = \int_{-\infty}^{\infty} |x| f(x) dx < \infty$$

1.3.2. Várható érték tulajdonságai

- Ha X az 1 valószínűséggel korlátos valószínűségi változó, akkor van olyan x_1 és x_2 konstans, hogy $P(x_1 \leq X \leq x_2) = 1$ akkor $x_1 \leq E(X) \leq x_2$
- $E(cX) = cE(X)$
- $P(X = c) = 1 \to E(X) = c$
- $E(X + Y) = E(X) + E(Y)$
- $E(X * Y) = E(X) * E(Y)$

1.4. Kovariancia függvény

$$\begin{aligned} R_X(u) &= \mathrm{cov}(X_t, X_{t-u}) \\ &= E[(X_t - E(X_t))(X_{t-u} - E(X_{t-u}))] \end{aligned}$$

Ez itt nem a kovarianciamátrixot fogja vissza adni, hanem az eltérés közötti összefüggést.

1.4.1. Kovariancia függvény tulajdonságai

- Additivitás: Ha X és Y véletlen változók és a és b valós számok, akkor a kovarianciafüggvény additív, azaz:

  $$cov(aX + bY, Z) = a_cov(X, Z) + b_cov(Y, Z)$$

- Szimmetria: A kovarianciafüggvény szimmetrikus, azaz

  $$cov(X, Y) = cov(Y, X)$$

- Állandóság: Ha X és Y véletlen változók és a és b konstansok, akkor a kovarianciafüggvény állandó marad, ha mindkét változót a-val és b-vel eltoljuk. Azaz,

  $$cov(X + a, Y + b) = cov(X, Y)$$

- Nemnegativitás: A kovarianciafüggvény mindig nemnegatív, azaz

  $$cov(X, X) \geq 0$$

  Ha a két változó független, akkor az egyenlőség akkor és csak akkor áll fenn, ha az X állandó.

- Normálás: Ha X és Y normális eloszlásúak, akkor a kovarianciafüggvény teljesen meghatározza a két változó közötti kapcsolatot.
- Két független változó kovarianciája nulla: Ha X és Y független változók, akkor a kovarianciafüggvényük zérus:

$$cov(X, Y) = 0$$

1.5. Gauss folyamat

Egy folyamatot Gauss folyamatnak nevezünk, ha a következő tulajdonságokkal rendelkezik:

- Az összes véges dimenziós eloszlása Gauss-eloszlású kell legyen. Ez azt jelenti, hogy az összes véges dimenziós eloszlásfüggvény szimmetrikus, és a karakterisztikus függvénye exponenciális alakú kell legyen.
  - Korreláció mátrixokat mind meg kell nézni, hogy pozitívak-e.
- Az összes időpillanatra vonatkozó középérték és szórás azonos kell legyen. A folyamat homogénnek tekinthető.
  - Ezt homogenitás teszttel lehet ellenőrizni.
- Az összes időpillanatban értékeket vesz fel végtelen dimenziós vektorokban. A végtelen dimenziós eloszlás azonban nem kell Gauss-eloszlásúnak lennie.

1.6. Herglotz-tétel

Legyen $R_X(u)$ a folyamat kovarianciafüggvénye, és tegyük fel, hogy ez a függvény az időbeli eltolásra invariáns, azaz csak a két időpont közötti különbségtől függ. Ekkor $R_X(u)$ Herglotz-féle sűrűségfüggvényként is felírható, azaz teljesül rá a következő:

$$R_X(u) = \int_{-\infty}^{\infty} e^{i*\lambda*u} * g_X(\lambda) d\lambda$$

ahol $g_X(\lambda)$ egy valós, szigorúan monoton növekvő eloszlásfüggvény. Más szóval, a kovarianciafüggvény Fourier-transzformáltját egy valós eloszlásfüggvénnyel lehet leírni.

2. Stacionárius folyamatok

A stacionárius folyamatok olyan valószínűségi folyamatok, amelyeknek a statisztikai tulajdonságai nem változnak az idő múlásával. Az ilyen folyamatok esetében a várható érték és a kovariancia függvénye nem függ az időtől, vagyis az idősor jellege nem változik az idő múlásával.

A stacionárius folyamatok matematikailag jól definiáltak és számos fontos tulajdonsággal rendelkeznek, amelyek lehetővé teszik számunkra az idősorok modellezését és előrejelzését. Az ilyen folyamatokra vonatkozóan meghatározott várható érték és kovariancia függvény jellemzi a folyamatot teljes egészében.

A stacionárius folyamatok fontosak a való életben előforduló idősorok modellezésében is, például a gazdasági mutatók és a meteorológiai adatok előrejelzésében. Az ilyen folyamatok matematikai tulajdonságai lehetővé teszik az idősorok előrejelzését, a kockázatbecslést és az optimalizálást.

2.1. Tágabb értelemben stacionárius folyamat

Legyen $\{X_t, t \in \mathcal{T}\}$, ahol

- X a t időponthoz tartozó sztochasztikus folyamat

A tágabb értelemben vett stacionárius folyamatot szokás röviden stacionárius folyamatnak nevezni.

Ha egy sztochasztikus folyamat tágabb értelemben stacionárius, az azt jelenti, hogy a várható értéke és a kovariancia függvénye csak az időbeli különbségtől függ, és nem az abszolút időponttól.

2.1.1. Tesztelés tágabb értelemben vett stacionárius folyamatra

Amennyiben a következő feltételek megegyeznek az egy tágabb értelemben vett stacionárius folyamat:

- $\mu_X(t) = E(X_t)$
- $R_X(s,t) = cov(X_s, X_t) = E(X_t - \mu_X(t))(X_s - \mu_X(s))$

Ahol a következők a következőket jelenti:

- $t, s \in \mathcal{T}$, tehát időbéli változók

- $\mu_X(t)$ egy konstans, ami csak az időtől függ és megegyezik a várható értékkel
- R_X a kovariancia függvény
- $E(X_t^2) < \infty$

2.2. Szűkebb értelemben stacionárius folyamat

Ahhoz, hogy valamit szűkebb értelemben stacionáriusnak nevezzünk teljesülnie kell, hogy $(X_{t_1}, \ldots, X_{t_n})$ és $(X_{t_{1+t}}, \ldots, X_{t_n+t})$ valószínűségi változók együttes eloszlása megegyezik és tágabb értelemben stacionárius folyamat.

3. Stacionárius Gauss folyamatok

A stacionárius Gauss-folyamat olyan Gauss-folyamat, amelynek a statisztikai tulajdonságai (középérték, szórás, autokorrelációs függvény) időtől függetlenek, vagyis az időbeli változások nem befolyásolják ezeket a tulajdonságokat.

4. Spektrál előállítás

A Herglotz-tétel szerint, a kovarianciafüggvényt kitudjuk fejezni a következő képen:

$$R_X(u) = \int_{-\infty}^{\infty} e^{i*\lambda*u} * g_X(\lambda) d\lambda$$

Ebben az összefüggésben a spektrális sűrűségfüggvény a $g_X(\lambda)$.

Analóg módon értelmezzük a diszkrét esetet is:

$$R_X(u) = \sum_{k=-\infty}^{\infty} \sigma_k^2 e^{i*u*\lambda_k}$$

4.1. Spektrális sűrűségfüggvény előállítása speciális esetben

Ha $\sum_{k=-\infty}^{\infty} |R_X(u)| < \infty$ feltétel teljesül, akkor a spektrális sűrűségfüggvény közvetlenül is előállítható:

$$g_X(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} R_X(k) e^{-i*k*\lambda}$$

Ez tovább írható:

$$g_X(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} R_X(k) cos(k\lambda) =$$
$$= \frac{1}{2\pi} \left(\sigma_X^2 + 2 \sum_{k=1}^{\infty} R_X(k) cos(k\lambda) \right)$$

4.2. Spektrális sűrűségfüggvény segítségével lévő stacionárius folyamat előállítása

$$X_t = \mu_X + \int_{-\pi}^{\pi} e^{it\lambda} dZ(\lambda)$$

$$X_t = \mu_X + \sum_{k=-\infty}^{\infty} Z_k e^{i*t*\lambda_k}$$

- μ_X a várható érték
- t az idő változó
- $Z(\lambda), -\pi \leq \lambda \leq \pi$ egy sztochasztikus folyamat, amely zérus várható értékű
- $E(Z(\lambda'' - Z(\lambda'))^2 = G_X(\lambda'') - G_X(\lambda')$, ha $-\pi \leq \lambda' < \lambda'' \leq \pi$

5. Becslések

5.1. Várható érték becslése

Legyen a várható érték becslése:

$$\bar{X} = \frac{1}{T} \sum_{i=1}^{T} X_i$$

- T az összes idő megfigyelése
- X a sztochasztikus folyamat.

Mint minden más statisztikai becsléstől ettől is elvárjuk a torzítatlanságot.

5.1.1. Torzítatlansága a várható érték

$$E(\bar{X}) = \frac{1}{T}\sum_{i=1}^{T} X_i = \frac{1}{T}\sum_{i=1}^{T}\mu_X = \mu_X$$

Így látszódik, hogy torzítatlan.

5.2. Kovariancia függvény becslése

A becsléshez használt összefüggések a következőek:

$$\hat{R}_{1,k} = \frac{1}{T}\sum_{j=1}^{T-|k|}(X_j - \mu_X)(X_{j+|k|} - \mu_X)$$

$$\bar{R}_{1,k} = \frac{1}{T-|k|}\sum_{j=1}^{T-|k|}(X_j - \mu_X)(X_{j+|k|} - \mu_X)$$

Ahol:

- T hosszú folyamatunk van
- k a "késeltetés" két megfigyelés között
- μ_X a várható érték, ha ez nem ismert érdemes becsülni.

6. Fehérzaj folyamat

Olyan stacionárius folyamat, amely

- korrelálatlan sorozatot alkot
- várható értéke 0
- Minden időpillanatban megegyezik az eloszlása.

6.1. Fehérzaj folyamat spektrális sűrűségfüggvénye

$$g_\epsilon(\lambda) = \frac{1}{2*\pi}\sum_{j=-\infty}^{\infty} R_\epsilon(j)e^{-ij\lambda}$$

6.2. standard fehérzaj folyamat

Olyan fehérzaj folyamat, amelynek a varianciája minden időpontra pontosan 1.

6.2.1. standard fehérzaj folyamat spektrális sűrűségfüggvénye

$$g_\epsilon(\lambda) = \frac{1}{2\pi}\sigma_\epsilon^2, \quad -\pi \leq \lambda \leq \pi$$

7. Harmonikus folyamatok

$$X_t = A_0 + \sum_{i=1}^{q}(A_i cos(\lambda_i t) + B_i sin(\lambda_i t))$$

- X_t-t nevezzük harmonikus folyamatnak, ha
- A_i és B_i korrelálatlan [[valószínűségi változó—valószínűségi változók]]
- A_i és B_i [[várható érték—várható értékei]] 0-k
- $VAR(A_0) = \sigma_0^2$
- $VAR(A_i) = VAR(B_i) = \sigma_i^2$
- $\lambda \in \mathbb{R}$, $q \in \mathbb{Z}^+$

7.1. Harmonikus folyamatok tulajdonságai

- Periodikusak
- csak a két időpont közötti távolságtól függ a két időpont korrelációja
- Gauss eloszlásúak
- nem létezik spektrális sűrűségfüggvénye

II. rész
Lineáris folyamatok

Elsőként szeretném szemléltetni a lineáris folyamatok használatát a következő példával:

Tegyük fel, hogy azt szeretnénk vizsgálni, hogy hogyan változnak az egyik tőzsdén jegyzett részvényáraink az idő függvényében. Az idősor analízis esetében a részvényárfolyamokat időpontonként mérjük és rögzítjük, és az időpontok közötti különbségek általában egyenlők.

Az ilyen típusú adatokat legegyszerűbb módon az autokorrelációs függvények segítségével modellezni. Az autokorrelációs függvény egy olyan matematikai eszköz, amely azt mutatja meg, hogy milyen erős az idősorbeli adatok közötti kapcsolat, azaz hogy az előző adatok milyen mértékben hatnak az időpontban mért adatokra.

Az autokorrelációs függvényekből pedig kiszámítható a spektrális sűrűségfüggvény is, amely az idősorbeli adatok frekvenciára vetített változását írja le. Ez azért fontos, mert az idősorbeli adatokban található frekvenciák meghatározzák az adatok jellemzőit, például azt, hogy milyen időtartamú ingadozások jellemzik az adott idősorbeli adatokat.

Ezeket a matematikai eszközöket alkalmazva az idősorbeli adatokat leíró lineáris folyamatokat lehet definiálni. Az ilyen lineáris folyamatok modelljei pedig felhasználhatók az idősorbeli adatok további vizsgálatára, például a trendek, szezonális változások vagy az idősorbeli adatok szabályszerűségeinek feltárására.

Egy ilyen példa egy lineáris folyamatra, amely az idősorbeli adatokat modellezi, lehet az ARIMA (AutoRegressive Integrated Moving Average) modell.

Azt a X_t folyamatot nevezzük lineáris folyamatnak, amely teljesíti a következő:

$$X_t = \sum_{s=-\infty}^{\infty} a_s \epsilon_{t-s}$$

Ahol:

- X_t egy sztochasztikus folyamat.
- $a_s \in \mathbb{R}$ és $\sum_{s=-\infty}^{\infty} a_s^2 < \infty$.
- $\forall \epsilon$ fehérzaj folyamat, $\sigma_\epsilon^2 > 0$.
- $\lim_{m,n\to\infty} E\left(X_t - \sum_{s=-m}^{n} a_s \epsilon_{t-s}\right)^2 = 0$.

8. Kauzális folyamat

A kauzális folyamat olyan lineáris folyamat, amely előállítható a következő képen:

$$X_t = \sum_{s=0}^{\infty} a_s \epsilon_{t-s}$$

Emellett az X_t csak a múltjától függ. Ezen felül egy lineáris folyamat akkor és csak akkor kauzális folyamat, ha a spektrális sűrűségfüggvénye létezik és teljesül rá a Bochner-Kolmogorov tétel:

$$\int_{-\pi}^{\pi} \ln(g_X(\lambda))d\lambda > -\infty$$

- g_X a spektrális sűrűségfüggvény
- λ időbeli változások sebességét leíró paraméter

8.1. Kauzális folyamat spektruma

- A kauzális folyamat spektruma kauszális: $g_X(\lambda) = 0$ minden $\lambda < 0$ esetén.
- $g_X(0) = \text{var}(X_t)$, azaz a spektrális sűrűségfüggvény értéke a zérus frekvencián megegyezik a folyamat varianciájával.
- A spektrális sűrűségfüggvény Fourier-transzformáltja pozitív mértékű véges mértékű eloszlás.
- Az autokovariancia függvény Fourier-transzformáltja az ún. spektrális eloszlásfüggvény, amelynek integrálja a folyamat varianciájával egyezik meg.
- Ha X_t kauzális és $Y_t = f(X_t)$, ahol f egy lineáris időinvariáns szűrő, akkor Y_t is kauzális és $g_Y(\lambda) = |f(\lambda)|^2 g_X(\lambda)$.
- A kauzális folyamatokat jellemezhetjük az ún. kauzalitási impulzusfüggvényükkel, amely az X_t folyamat egy egység-Dirac impulzus bemenetre adott válaszfüggvénye.

9. Lineáris szűrő

Lineáris szűrőt jelölje L.

$$L(X_t) = \sum_{s=-\infty}^{\infty} h(t-s)X_s$$

- Lineáris szűrőt jelölje L
- $L(X_t) = \sum_{s=-\infty}^{\infty} h(t-s)X_s$
- X_t egy vagy több dimenziós stacionárius folyamat (tágabb értelemben)
  - Van a folyamatnak $R_x(u)$-val definiált kovarianciafüggvénye
  - $g_X(\lambda)$ spektrális sűrűségfüggvény is létezik
  - $E(X_t) \equiv 0$
- $h(t)$ függvény eleget tesz a szűrő koherenciafeltételnek:

$$\sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} h(u)R_X(v-u)h^T(v) < \infty$$

9.1. transzfer függvény

A szűrő koherenciafeltételhez tartozik az úgynevezett transzfer függvény:

$$H\left(e^{-i\lambda}\right) = \sum_{t=-\infty}^{\infty} e^{-it\lambda}h(t)$$

- H a transzfer függvény
- λ az időbeli változások sebességét leíró paraméter.

Legyen $Y_t = L(X_t)$ stacionárius folyamat, ahol:

- L egy lineáris szűrő
- X_t egy stacionárius folyamat

Ekkor az Y_t kovarianciafüggvényére fennáll:

$$R_Y(r, r+t) = R_Y(t)$$

Ahol:

- r a vizsgált időpont
- t pedig, hogy mennyi időegységre van eltolva a vizsgálat vége az r-től

Az Y_t spektrális sűrűségfüggvényére pedig:

- Több dimenziós esetben a $g_Y(\lambda) = H\left(e^{-i\lambda}\right) g_X(\lambda) H^{T*}\left(e^{-i\lambda}\right)$ összefüggés igaz.
- Egy dimenziós esetben pedig a $g_Y(\lambda) = g_X(\lambda)\left|H\left(e^{-i\lambda}\right)\right|^2$ összefüggés igaz.

Ahol:

- g a spektrális sűrűségfüggvény
- $\square^{T*}$ a transzponált, majd konjugált műveletek jelölik.

$H\left(e^{-i\lambda}\right)$ esetén a kauzális folyamat spektrális sűrűségfüggvénye a

$$g_X(\lambda) = \frac{\sigma_\epsilon^2}{2\pi}\left|\sum_{t=0}^{\infty} a_t e^{-it\lambda}\right|^2$$

lesz.

10. autókorreláció függvény

Megmutatja, hogy a folyamat időbeli részei mennyire hasonlítanak egymáshoz, mennyire van közös mozgásuk. Legyen X_t stacionáriu] folyamat $R_X(t)$ pedig kovarianciafüggvény, ekkor az autokorrelációs függvényt a következő képen értelmezzük:

$$r_X(t) = \frac{1}{R_X(0)} R_X(t) = \frac{1}{\sigma_X^2} R_X(t)$$

.

10.1. Autókorreláció függvény spektruma

A következőt nevezzük az autokorrelációs függvény spektrumának:

$$g_X(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} r_X(k) e^{-i*k*\lambda}$$

A zaj, az interferencia vagy a periodikus mintázatok kezelésében. A spektrum azonban csak egy részleges képet ad a folyamatról, és egyedül nem elegendő a folyamat jellemzéséhez. Számos más jellemzőnek, mint például az autoregressziós modell paramétereinek, a szűrő paramétereinek vagy a spektrális sűrűségfüggvénynek ismerete szükséges lehet a folyamatok részletes elemzéséhez.

10.2. parciális autokorreláció

$Z_1, \ldots, Z_k$ valószínűségi változók melletti parciális korrelációnak nevezzük a következő összefüggést:

$$\rho = \mathrm{CORR}(X - \hat{X}, Y - \hat{Y}) = \frac{\mathrm{COV}(X - \hat{X}, Y - \hat{Y})}{D(X - \hat{X})D(Y - \hat{Y})}$$

Ahol:

- X, Y, Z valószínűségi változók
- $\hat{X} = \hat{X}(Z_1, \ldots, Z_k), \hat{Y} = \hat{Y}(Z_1, \ldots, Z_k)$
- $D(X - \hat{X}) > 0, D(Y - \hat{Y}) > 0.$

Megmutatja az X és Y közötti kapcsolat erősségét azután, hogy mindkét változóban kiküszöböltük a Z valószínűségi változók hatását.

10.2.1. parciális autokorreláció stacionárius folyamat esetén

Azt mutatja meg, hogy két időbeli pontra vonatkozó korreláció mennyiben magyarázható a közöttük lévő időbeli pontok hatásának kiküszöbölésével. Stacionárius folyamatok esetén a parciális autokorreláció különösen fontos, mivel lehetővé teszi számunkra, hogy azonosítsuk a folyamat autoregresszív (AR) modellének paramétereit. X_t és X_{t-k} közötti parciális autokorreláció alatt a következő összefüggést értjük:

$$\rho_k = CORR(X_t - \widehat{X}_t, X_{t-k} - \widehat{X}_t) = \frac{COV(X_t - \widehat{X}_t, X_{t-k} - \widehat{X}_t)}{D(X_t - \widehat{X}_t)D(X_{t-k} - \widehat{X}_t)}$$

ahol:

- $k \in \mathbb{Z}$,
- $\widehat{X}_t = \widehat{X}_t(X_{t-1}, \ldots, X_{t-k+1})$,
- $D(X_t - \hat{X}_t) > 0$ és $D(X_{t-k} - \hat{X}_t) > 0$.

11. AR; MA; ARMA folyamatok

Ide valami arról, hogy mire való

11.1. AR

Kibővíteni

11.2. MA

Ez mi

11.3. ARMA

A kettő együtt mit jelent

12. Paraméterek becslése

III. rész
Wiener folyamatok