Výpočet středních hodnot metodou Monte Carlo

Úkol
Generujte Markovovy řetězce bodů na reálné ose pro Gaussovy distribuce vybraných rozptylů a sledujte konvergenci různých populací těchto bodů k zadaným rovnovážným distribucím.

Pomocí získaných dat vypočtěte pro všechny volby počtu bodů v populaci a pro všechny volby rozptylů Gaussových distribucí střední hodnoty náhodné proměnné a hodnoty její střední kvadratické fluktuace a porovnejte je s hodnotami teoretickými.

Použitý software
Předkompilovaný program `u10.exe`, Origin

Použité soubory se zdrojovými kódy
V této úloze budete používat předkompilovaný program a se zdrojovými kódy pracovat nebudete.

Popis vstupních souborů
`main.ini`

`INTEGER           :: NMB_POINTS`	– počet generovaných bodů
`DOUBLE PRECISION :: SIGMA`	– rozptyl Gaussovy distribuce
`DOUBLE PRECISION :: X_INI`	– první bod Markovova řetězce
`DOUBLE PRECISION :: DX`	– maximální náhodný posun bodu v jednom kroku

Teorie
Normovaná ***Gaussova distribuce*** centrovaná v počátku reálné osy je dána předpisem

$$\rho(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}},$$

kde σ je tzv. rozptyl této distribuce. Pro ***střední hodnotu a střední kvadratickou fluktuaci*** náhodné proměnné x platí v případě Gaussovy distribuce centrované v počátku souřadnic

$$\langle x \rangle \equiv \int_{-\infty}^{+\infty} x\rho(x)\,\mathrm{d}x = 0$$

a

$$\Delta x^2 \equiv \langle x^2 \rangle - \langle x \rangle^2 = \sigma^2 .$$

Markovův řetězec $\{x^{(i)}\}_{i=1}^{N}$ generujeme na reálné ose pro zadanou rovnovážnou distribuci rekurentním dvoukrokovým algoritmem (i a i+1 indexují i-tý a (i+1)-ní člen řetězce)

1. vygenerujeme nový bod $x_{\mathrm{new}} \equiv x^{(i)} + R \cdot \delta x$, kde $x^{(i)}$ je bod vygenerovaný v předcházejícím kroku, δx konstantní parametr zadávající maximální posun bodu v jednom kroku a R náhodné číslo z intervalu (-1,1);

2. s pravděpodobností $\alpha \equiv \min\{1, \rho(x_{\text{new}}) / \rho(x^{(x_i)})\}$ položíme $x^{(i+1)} = x_{\text{new}}$ a s pravděpodobností $1-\alpha$ položíme $x^{(i+1)} = x^{(i)}$.

V limitě $N \to +\infty$ bude hustota rozložení na reálné ose pro takto vytvořené body úměrná použité distribuci ρ. Pro konečný počet bodů ($N < +\infty$) odpovídá histogram rozložení bodů na reálné ose (po náležitém normování) distribuci ρ jen přibližně, leč s rostoucím N odchylky mezi generovaným histogramem a vzorovou distribucí klesají k nule. Říkáme, že výsledky výpočtu konvergují k ρ. Abychom tuto konvergenci optimalizovali, snažíme se obvykle eliminovat korelace mezi dvěma po sobě následujícími body, $x^{(i+1)}$a $x^{(i)}$, které do výpočtu histogramu zahrnujeme. V našem programu tak činíme obvyklým způsobem a do histogramu (statistického vyhodnocení dat) zahrnujeme každý stý generovaný bod.

Normalizovanou hustotu rozložení bodů řetězce $\{x^{(i)}\}_{i=1}^{N}$ na reálné ose (přímo porovnatelnou s normalizovanou Gaussovou distribucí) aproximujeme vztahem

$$\eta(x) \approx \frac{\Delta N(x, x+\Delta x)}{N \Delta x},$$

kde Δx je dostatečně malý přírůstek náhodné proměnné a $\Delta N(x, x+\Delta x)$ počet bodů řetězce $\{x^{(i)}\}_{i=1}^{N}$ z intervalu $(x, x+\Delta x)$ (tj. nenormovaný histogram).

Postup práce

1. Proveďte výpočty pro různá nastavení souboru `main.ini` (po modifikaci nezapomeňte soubor `main.ini` vždy uložit; pro spuštění poklepejte na program `u10.exe`). V pracovní složce po každém výpočtu vhodně přejmenujte výstupní soubor `uloha10.txt` (např. tak, jak navrhujeme níže).
   - `SIGMA` = 5, `DX` = 0.5, `NMB_POINTS` = 100 (výstupní soubor `sg50nmb1E2.txt`),
   - `SIGMA` = 5, `DX` = 0.5, `NMB_POINTS` = 1000 (výstupní soubor `sg50nmb1E3.txt`),
   - `SIGMA` = 5, `DX` = 0.5, `NMB_POINTS` = 10000 (výstupní soubor `sg50nmb1E4.txt`),
   - `SIGMA` = 5, `DX` = 0.5, `NMB_POINTS` = 50000 (výstupní soubor `sg50nmb5E4.txt`),
   - `SIGMA` = 1, `DX` = 0.1, `NMB_POINTS` = 100 (atd.),
   - `SIGMA` = 1, `DX` = 0.1, `NMB_POINTS` = 1000,
   - `SIGMA` = 1, `DX` = 0.1, `NMB_POINTS` = 10000,
   - `SIGMA` = 1, `DX` = 0.1, `NMB_POINTS` = 50000,
   - `SIGMA` = 0.1, `DX` = 0.01, `NMB_POINTS` = 100,
   - `SIGMA` = 0.1, `DX` = 0.01, `NMB_POINTS` = 1000,
   - `SIGMA` = 0.1, `DX` = 0.01, `NMB_POINTS` = 10000,
   - `SIGMA` = 0.1, `DX` = 0.01, `NMB_POINTS` = 50000.
2. Spusťte program Origin a načtěte výstupní textové soubory s vygenerovanými Markovovými řetězci (opakovaně zvolte `File/Import/Single ASCII` a vyberte požadovaný soubor).
3. Pomocí `Statistics/Descriptive Statistics/Frequency Count` vytvořte pro každý řetězec tabulku s rozložením bodů na reálné ose s následujícími nastaveními v dialogovém boxu
   - `From Minimum` = -2 `SIGMA`,
   - `To Maximum` = 2 `SIGMA`,
   - `Step Size` = `SIGMA` /10.
4. Každou nově vytvořenou tabulku (worksheet) vhodně pojmenujte a sloupec `Count` normujte označením, volbou `Column/Set Column Values` a zadáním v dialogovém boxu `col(Count)/Step/Nmb`, kde `Step` je číselná hodnota `Step Size` a `Nmb` číselná hodnota `NMB_POINTS`.

5. Sloupce `Count` vytvořené v kroku (3) vykreslete do grafu (označte sloupec a klikněte na ikonu sloupcového grafu) a porovnejte s odpovídající distribucí (zvolte `Graph/Add Function Graph` a do dialogoveho boxu zadejte `1/SIGMA/(6.2832)^(0.5)*exp(-x^2/2/SIGMA^2)` , kde `SIGMA` je číselná hodnota parametru).
6. Pro každou datovou sadu z kroku (1) určete střední hodnoty a střední kvadratické fluktuace náhodné veličiny *x*, shrňte je do přehledné tabulky a porovnejte s přesnými hodnotami odpovídajícími parametrům použité Gaussovy distribuce. (Výpočet provádíte vždy v konkrétní tabulce vytvořené v kroku (1) označením sloupce s Markovovým řetězcem a volbou `Statistics/Descriptive Statistics/Statistics on Columns`. Data pak odečítáte ze sloupců `Mean` a `sd` nově vytvořené tabulky, kterou po odečtení smažete.)

Doporučená literatura

literatura k lekci 11 kurzu KFY/PMFCH (viz http://artemis.osu.cz/pmfch/lekce11.pps nebo http://artemis.osu.cz/pmfch/lekce11.pdf)
manuál k software Origin