Метод 2D-SSA для анализа двумерных полей

Усевич Константин Дмитриевич, гр. 522

Санкт-Петербургский государственный университет
Математико-механический факультет
Кафедра статистического моделирования

Научный руководитель: к.ф.-м.н. Некруткин В.В.
Рецензент: к.ф.-м.н. Голяндина Н.Э.

СтатМод

Санкт-Петербург
2007г.

Анализ двумерных дискретных полей

- двумерное поле – функция двух переменных $f(x, y)$
- дискретное – функция задана на дискретной равномерной сетке

Задача: разложить поле $f(x, y)$ на составляющие в соответствии с внутренней структурой.

Методы SSA (SSA,MSSA,2D-SSA):

- Функция $f : D \mapsto \mathbb{R}$, D – дискретное пространство
- Параметры $\mathfrak{T}$ – размеры окна.

1. Разложение
   - Вложение $f \rightsquigarrow \mathbf{W}$ (инъекция)
   - Сингулярное разложение
     $\mathbf{W} = \sum_{i=1}^{d} \sqrt{\lambda_i} U_i {V_i}^{\mathrm{T}}$

SSA, ряд $(f_0, \ldots, f_{N-1})$

$$\mathbf{W} = \begin{pmatrix} f_0 & f_1 & f_2 & \cdots & f_{K-1} \\ f_1 & f_2 & f_3 & \cdots & f_K \\ f_2 & f_3 & \cdots\cdots & \cdots & f_{K+1} \\ \vdots & \vdots & \cdots\cdots & \cdots & \vdots \\ f_{L-1} & f_L & \cdots\cdots & \cdots & f_{N-1} \end{pmatrix}$$

2. Восстановление
   - Группировка $\mathbf{W} = \mathbf{W}_{I_1} + \cdots + \mathbf{W}_{I_m}$, где $\mathbf{W}_I = \sum_{i \in I} \sqrt{\lambda_i} U_i {V_i}^{\mathrm{T}}$
   - Ортогональное проектирование $\mathbf{W}_I \dashrightarrow f_I$
     $f = f_{I_1} + \cdots + f_{I_m}$

План дальнейшего изложения

Непрерывный случай

- Методы SSA на функциональном языке
- Непрерывный 2D-SSA
- Результаты для абстрактного и непрерывного случая

Дискретный случай

- Алгоритм
- Разделимость
- Поля конечного ранга
- Отделение от шума – результаты моделирования
- Программа, пример
- Приложения 2D-SSA

SSA-методы на функциональном языке

Рассматривается разложение измеримой функции $f : (D, \mathcal{U}) \mapsto \mathbb{R}$.

- Измеримые пр-ва $(D_1, \mathcal{U}_1, \mu_1)$, $(D_2, \mathcal{U}_2, \mu_2)$; μ_1, μ_2 – конечные меры.
- Пространства функций

$$\mathbf{L}_1^2 = \mathbf{L}^2(D_1, d\mu_1), \mathbf{L}_2^2 = \mathbf{L}^2(D_2, d\mu_2), \mathbf{L}_{1,2}^2 = \mathbf{L}^2(D_1 \times D_2, d(\mu_1 \otimes \mu_2)),$$

- Измеримая сюръекция $\theta : D_1 \times D_2 \mapsto D$.
- θ порождает меру ν на $(D, \mathcal{U})$ и σ-алгебру $\mathcal{U}_\theta$ на $D_1 \times D_2$:

$$\nu(A) = \mu_1 \otimes \mu_2(\theta^{-1} A),\ A \in \mathcal{U}; \quad \mathcal{U}_\theta = \{\theta^{-1}(A), A \in \mathcal{U}\}.$$

- Для $f \in \mathbf{L}^2(D, d\nu)$ существует разложение Шмидта $g = f \circ \theta \in \mathbf{L}_{1,2}^2$

$$g(u, s) = f(\theta(u, s)) = \sum\nolimits_n \sqrt{\lambda_n}\, \psi_n(u)\, \varphi_n(s), \quad \psi_n \in \mathbf{L}_1^2, \varphi_n \in \mathbf{L}_2^2,$$

- Ортогональные проекции $\mathfrak{P}_\theta : \mathbf{L}_{1,2}^2 \mapsto \mathbf{L}^2(D_1 \times D_2, \mathcal{U}_\theta, d(\mu_1 \otimes \mu_2|_{\mathcal{U}_\theta}))$

$$f_n \circ \theta = \mathfrak{P}_\theta(\sqrt{\lambda_n} \psi_n \otimes \varphi_n).$$

- Переход к разложению $f = \sum\nolimits_n f_n$

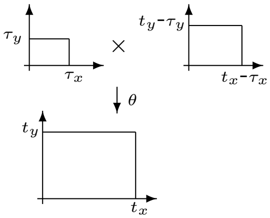
Непрерывный 2D-SSA

Разложение функции $f : [0, t_x) \times [0, t_y) \mapsto \mathbb{R}$.

- Параметры: размеры окна (τ_x, τ_y)

$$\begin{aligned} D_1 &= [0, \tau_x) \times [0, \tau_y), \\ D_2 &= [0, t_x - \tau_x] \times [0, t_y - \tau_y]. \end{aligned}$$

- $\theta((u_x, u_y), (s_x, s_y)) = (u_x + s_x, u_y + s_y)$.



- Меры μ_1, μ_2 – меры Лебега (равномерные распределения).
- Мера $\nu = \mu_1 \otimes \mu_2 \circ \theta^{-1}$ – сумма равномерных распределений.

Алгоритм:

1. Этап разложения
   - Вложение $g\big((u_x, u_y), (s_x, s_y)\big) = f(u_x + s_x, u_y + s_y)$
   - Разложение Ш. $g\big((u_x, u_y), (s_x, s_y)\big) = \sum_n \sqrt{\lambda_n}\, \psi_n(u_x, u_y)\, \varphi_n(s_x, s_y)$
2. Этап восстановления
   - Группировка $g = g_{I_1} + \cdots + g_{I_m}$, где $g_I = \sum_{i \in I} \sqrt{\lambda_i}\, \psi_i \otimes \varphi_i$
   - Проектирование $f = f_{I_1} + \cdots + f_{I_m}$, где $f_I = \mathfrak{R}_\theta(g_I)$

Результаты для непрерывного (абстрактного) случая

Разложение Шмидта

$$f(u_x + s_x, u_y + s_y) = \sum_n \sqrt{\lambda_n}\, \psi_n(u_x, u_y)\, \varphi_n(s_x, s_y)$$

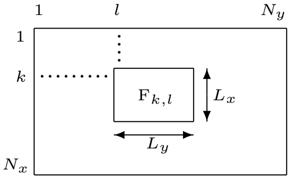
- Сформулировано понятие разделимости, выведены условия разделимости в непрерывном случае.
- Для функций вида $f(x,y) = p(x)q(y)$:
  - Вид разложения через разложение одномерных функций
  - Условие разделимости $f_1(x,y) = p_1(x)q_1(y)$ и $f_2(x,y) = p_2(x)q_2(y)$
- Введено понятие поля конечного ранга (с конечным числом слагаемых в разложении). Доказаны утверждения для вычисления ранга, в т.ч. для конкретных полей: полиномов.
- Доказана инвариантность в смысле сохранения ранга относительно линейных преобразований
  - одномерных функций $p(ax + by)$
  - полей $f(a_1x + b_1y, a_2x + b_2y)$

Алгоритм дискретного 2D-SSA, Этап 1: разложение

Разложение дискретных двумерных полей

- Поле F размера $N_x \times N_y$

$$\mathrm{F} = \begin{pmatrix} f_{0,0} & \dots & f_{0,N_y-1} \\ \vdots & \ddots & \vdots \\ f_{N_x-1,0} & \dots & f_{N_x-1,N_y-1} \end{pmatrix}$$



- Параметры: размеры окна (L_x, L_y)

Из подматриц $L_x \times L_y$ строится матрица $\mathbf{X}$

$$\mathrm{F}_{k,l} = \begin{pmatrix} f_{k-1,l-1} & \dots & f_{k-1,l+L_y-2} \\ \vdots & \ddots & \vdots \\ f_{k+L_x-2,l-1} & \dots & f_{k+L_x-2,l+L_y-2} \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} \mathrm{F}_{1,1} & \dots & \mathrm{F}_{1,K_y} \\ \vdots & \ddots & \vdots \\ \mathrm{F}_{K_x,1} & \dots & \mathrm{F}_{K_x,K_y} \end{pmatrix}$$

Разложение $\mathbf{X}$

$$\mathbf{X} = \sum_{i=1}^{m} \mathrm{Q}_i \otimes \mathrm{P}_i,$$
$$\mathrm{P}_i \in \mathcal{M}_{L_x,L_y}, \mathrm{Q}_i \in \mathcal{M}_{K_x,K_y}$$

Кронекеровское произведение

$$A \otimes B \stackrel{\mathrm{def}}{=} \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix}$$

Алгоритм дискретного 2D-SSA, Этап 1: разложение

Двумерная траекторная матрица $\mathbf{X}$

$$\mathbf{X} = \begin{pmatrix} \mathrm{F}_{1,1} & \mathrm{F}_{1,2} & \dots & \mathrm{F}_{1,K_y} \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{F}_{K_x,1} & \mathrm{F}_{K_x,2} & \dots & \mathrm{F}_{K_x,K_y} \end{pmatrix}$$

Векторизация

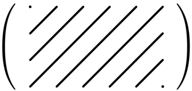
$$\operatorname{vec}\begin{pmatrix} \blacktriangle & \blacksquare & \dots & \star \\ \blacktriangle & \blacksquare & \dots & \star \\ \blacktriangle & \blacksquare & \dots & \star \end{pmatrix} \stackrel{\mathrm{def}}{=} = (\blacktriangle\ \blacktriangle\ \blacktriangle\ \blacksquare\ \blacksquare\ \blacksquare\ \dots\ \star\ \star\ \star)^{\mathrm{T}}$$

Матрица $\mathbf{W}$ из векторизованных подматриц $\quad W_{k+(l-1)L_x} = \operatorname{vec}(\mathrm{F}_{k,l})$

Блочно-ганкелева траекторная матрица $\mathbf{W}$

$$\mathbf{W} = \begin{pmatrix} \mathbf{H}_0 & \mathbf{H}_1 & \dots & \mathbf{H}_{K_y-1} \\ \mathbf{H}_1 & \mathbf{H}_2 & \dots & \mathbf{H}_{K_y} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_{L_y-1} & \mathbf{H}_{L_y} & \dots & \mathbf{H}_{N_y-1} \end{pmatrix},$$

$\mathbf{H}_j$ – ганкелевы блоки $L_x \times K_x$,

$\mathbf{H}_j =$ 

$\mathbf{H}_j$ – траекторные матрицы рядов $f(\cdot, j)$.

Разложение

KP-SVD $\quad \mathbf{X} = \sum_{i=1}^{d} \mathbf{X}_i = \sum_{i=1}^{d} \sqrt{\lambda_i}\Phi_i \otimes \Psi_i,$

SVD $\quad \mathbf{W} = \sum_{i=1}^{d} \mathbf{W}_i = \sum_{i=1}^{d} \sqrt{\lambda_i} U_i {V_i}^{\mathrm{T}},$

собственные	факторные
$U_i = \operatorname{vec}\Psi_i$	$V_i = \operatorname{vec}\Phi_i$

Разделимость

Поле является суммой $\mathrm{F} = \mathrm{F}^{(1)} + \mathrm{F}^{(2)}$

$$\mathbf{X} = \sum_{i=1}^{d} \sqrt{\lambda_i}\Phi_i \otimes \Psi_i \stackrel{?}{=} \overbrace{\sum_{i\in I_1} \sqrt{\lambda_i}\Phi_i \otimes \Psi_i}^{\mathrm{F}^{(1)}} + \overbrace{\sum_{i\in I_2} \sqrt{\lambda_i}\Phi_i \otimes \Psi_i}^{\mathrm{F}^{(2)}}$$

Определение

Траекторное пространство $\mathcal{L}^{(L_x,L_y)}(\mathrm{F}) \stackrel{\mathrm{def}}{=} \mathrm{span}\left(\{\mathrm{F}_{k,l}\}_{k,l=1}^{K_x,K_y}\right)$.

Поля $\mathrm{F}^{(1)}$ и $\mathrm{F}^{(2)}$ *слабо разделимы окном* (L_x, L_y), если $\mathcal{L}^{(L_x,L_y,1)} \perp \mathcal{L}^{(L_x,L_y,2)}$ и $\mathcal{L}^{(K_x,K_y,1)} \perp \mathcal{L}^{(K_x,K_y,2)}$.

Свойства

- $\mathrm{F}^{(1)}$ и $\mathrm{F}^{(2)}$, $\mathrm{F}^{(1)}$ и $\mathrm{F}^{(3)}$ – разделимы $\Rightarrow$ $\mathrm{F}^{(1)}$ и $\mathrm{F}^{(2)} + \mathrm{F}^{(3)}$ – разделимы
- Разделимость $\mathrm{F}^{(1)} = P_1 Q_1^{\mathrm{T}}$ и $\mathrm{F}^{(2)} = P_2 Q_2^{\mathrm{T}}$

Пример $f(i,j) = \ln(i+1)\sin(aj) + \sin(bi)\ln(j+1)$ отделимо от константы

Поля конечного ранга

Определение

(L_x, L_y)-*Ранг поля:* $\mathrm{rank}_{L_x,L_y}(\mathrm{F}) \stackrel{\mathrm{def}}{=} \dim \mathcal{L}^{(L_x,L_y)} = d.$

- Произведение рядов $(f(i,j) = p(i)q(j))$

$$\mathrm{rank}_{L_x,L_y}(F_x F_y^{\mathrm{T}}) = \mathrm{rank}_{L_x}(F_x)\,\mathrm{rank}_{L_y}(F_y)$$

- Полиномы

$$P_m(i,j) = \sum_{n=0}^{m}\sum_{k=0}^{n} a_{nk} i^k j^{n-k}$$

$$m+1 \;\le\; \mathrm{rank}_{L_x,L_y} P_m \;\le\; \begin{cases} \left(\lfloor \frac{m}{2} \rfloor + 1\right)\left(\lfloor \frac{m}{2} \rfloor + 1\right), & \text{для четных } m, \\ \left(\lfloor \frac{m}{2} \rfloor + 1\right)\left(\lfloor \frac{m}{2} \rfloor + 2\right), & \text{для нечетных } m. \end{cases}$$

- Тригонометрические функции

$$h_d(i,j) = \sum_{m=1}^{d} \begin{pmatrix} \cos(2\pi\omega_m^{(X)} i) & \sin(2\pi\omega_m^{(X)} i) \end{pmatrix} \begin{pmatrix} a_m & b_m \\ c_m & d_m \end{pmatrix} \begin{pmatrix} \cos(2\pi\omega_m^{(Y)} j) \\ \sin(2\pi\omega_m^{(Y)} j) \end{pmatrix}$$

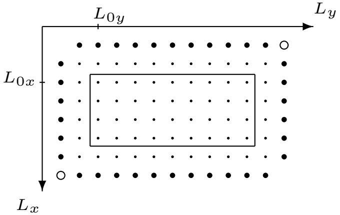
Поля конечного ранга

Определение

Будем говорить, что F – *поле конечного ранга*, если $\exists\, d, L_{x0}, L_{y0}$:

$$1 \le L_{x0} < \lfloor N_x/2 \rfloor,\ 1 \le L_{y0} < \lfloor N_y/2 \rfloor,\ 1 \le d \le L_{x0} \cdot L_{y0},$$

и для любых (L_x, L_y), т.ч. $L_{x0} \le \min(L_x, K_x)$, $L_{y0} \le \min(L_y, K_y)$, имеет место равенство $\mathrm{rank}_{L_x, L_y}(\mathrm{F}) = d$.



- Конечный ранг
  $\mathrm{rank}_{L_x, L_y}(\mathrm{F}) \equiv d$
- Неполный ранг
  $\mathrm{rank}_{L_x, L_y}(\mathrm{F}) < \min(L_x L_y, K_x K_y)$

Неполный ранг $\not\Rightarrow$ Конечный ранг

Пример

Ранг поля $f(i,j) = \ln(i+1)\sin(aj)$ равен $\min(L_x, K_x) \cdot 2$

Моделирование, отделение полей от шума

Сравнивались следующие методы (частные случаи 2D-SSA)

- SVD – размеры окна $(N_x, 1)$ или $(1, N_y)$

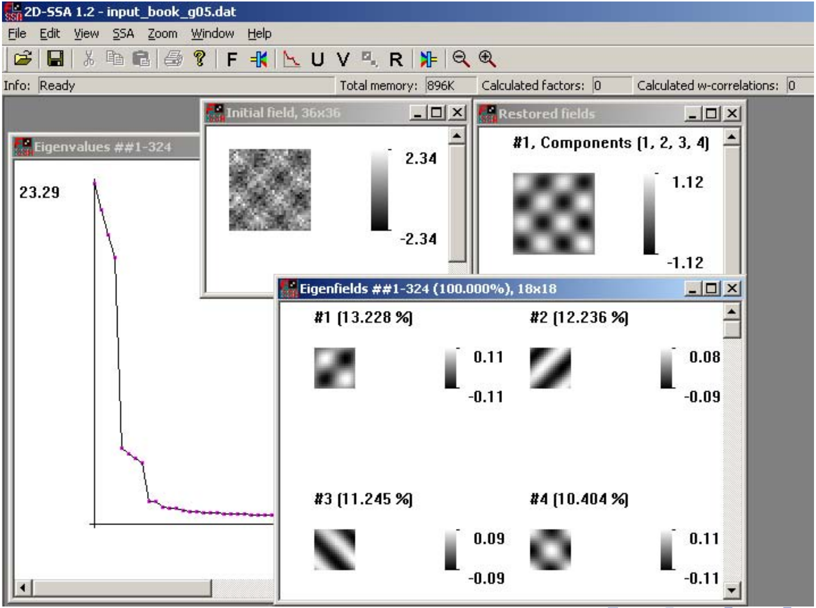
$$\mathrm{F} = \mathbf{W} = \sum_{i=1}^{d} \sqrt{\lambda_i} U_i {V_i}^{\mathrm{T}}$$

- MSSA – размеры окна $(L_x, 1)$ (MSSA-X) и $(1, L_y)$ (MSSA-Y)
- 2D-SSA – общий случай
- 2D-SSA – эквивалентный SVD, размеры окна (L_0, L_0)

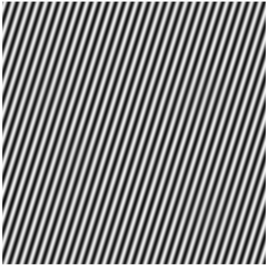
Результаты для полей

- конечного ранга
  - Восстанавливать лучше по рангу
  - 2D-SSA $(\lfloor N_x/2 \rfloor, \lfloor N_y/2 \rfloor)$ – 2D-SSA (L_0, L_0) – MSSA
  - Если ранг у методов одинаковый, то 2D-SSA (L_0, L_0) лучше SVD
- неполного ранга
  - SVD иногда лучше 2D-SSA (в случае $f(i,j) = \sum\limits_k p_k(i) q_k(j)$).
  - 2D-SSA (L_0, L_0) лучше 2D-SSA $(\lfloor N_x/2 \rfloor, \lfloor N_y/2 \rfloor)$
- произвольных изображений

Программа


2D-SSA 1.2 - input_book_g05.dat
File Edit View SSA Zoom Window Help
F U V R
Info: Ready
Total memory: 896K
Calculated factors: 0
Calculated w-correlations: 0
Initial field, 36x36
2.34
-2.34
Restored fields
#1, Components (1, 2, 3, 4)
1.12
-1.12
Eigenvalues ##1-324
23.29
Eigenfields ##1-324 (100.000%), 18x18
#1 (13.228 %)
0.11
-0.11
#2 (12.236 %)
0.08
-0.09
#3 (11.245 %)
0.09
-0.09
#4 (10.404 %)
0.11
-0.11

Пример: удаление периодичного шума



Пример: удаление периодичного шума