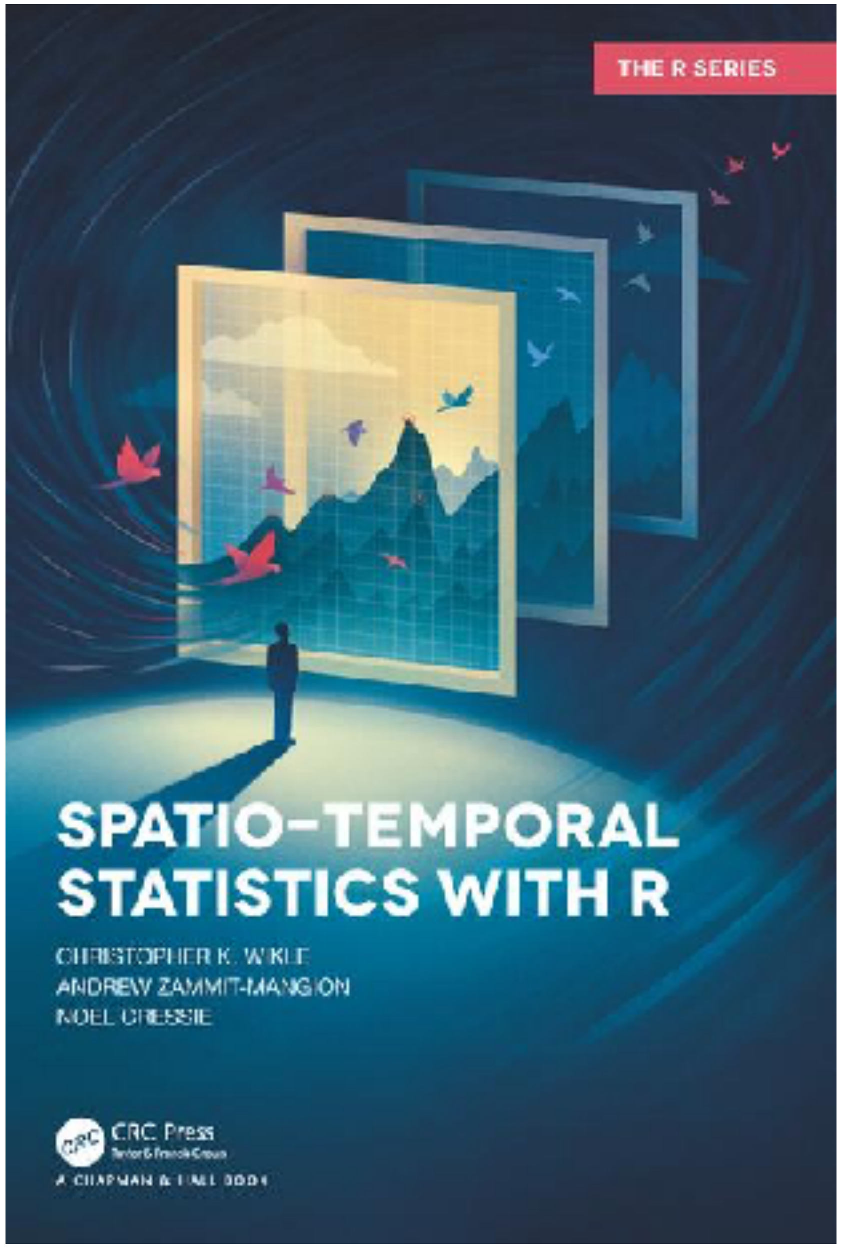


2.4 Exploratory analysis of spatio-temporal data

2.5 Chapter 2 wrap-up

Lab 2.3: Exploratory analysis of spatio-temporal data

2.4 Exploratory Analysis of Spatio-Temporal Data (p.32-)

- 一次・二次特徴の時空間データの要約
  - (2.4.1) empirical means, empirical covariance
  - (2.4.2) spatio-teoporal covariogramus and semivariograms
  - (2.4.3) empirical orthogonal functions and principal-component time series
  - (2.4.4) spatio-temporal canonical correlation analysis

- 数式の定義 -> Appendix A(p.307-)

A Some Useful Matrix-Algebra Definitions and Properties

For the sake of completeness, we provide some definitions and properties of vectors and matrices that are needed to understand many of the formulas and equations in this book. Readers who are already familiar with matrix algebra can skip this section. Readers who would like more detail than the bare minimum presented here can find them in books on matrix algebra or multivariate statistics (e.g., Johnson and Wichern, 1992; Schott, 2017).

Vectors and matrices. In this book we denote a *vector* (a column of numbers) by a bold letter (Latin or Greek); for example,

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix}$$

represents a p-dimensional vector, and $\mathbf{a}' = [a_1, a_2, \ldots, a_p]$ or $(a_1, a_2, \ldots, a_p)$ is its p-dimensional transpose.

We also denote a *matrix* (an array of numbers) by bold upper-case letters (Latin or Greek); for example,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix}$$

is a $p \times n$ matrix, and $a_{k\ell}$ corresponds to the element in the kth row and ℓth column; sometimes it is also written as $\{a_{k\ell}\}$. The matrix transpose, $\mathbf{A}'$, is then an $n \times p$ matrix given by

$$\mathbf{A}' = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{p1} \\ a_{12} & a_{22} & \cdots & a_{p2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{pn} \end{bmatrix}.$$

We often consider a special matrix known as the *identity matrix*, denoted $\mathbf{I}_n$, which is an $n \times n$ diagonal matrix with ones along the main diagonal (i.e., $a_{ii} = 1$ for $i = 1, \ldots, n$) and zeros for all of the off-diagonal elements (i.e., $a_{ij} = 0$, for $i \neq j$). It is sometimes the case that the dimensional subscript (in this case, n) is left off if the context is clear.

Finally, note that a vector can be thought of as a special case of a $p \times n$ matrix, where either $p = 1$ or $n = 1$.

Matrix addition. Matrix addition is defined for two matrices that have the same dimension. Then, given $p \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$, with elements $\{a_{k\ell}\}$ and $\{b_{k\ell}\}$ for $k = 1, \ldots, p$ and $\ell = 1, \ldots, n$, respectively, the elements of the matrix sum, $\mathbf{C} = \{c_{k\ell}\} = \mathbf{A} + \mathbf{B}$, are given by

$$c_{k\ell} = a_{k\ell} + b_{k\ell}, \quad k = 1, \ldots, p;\ \ell = 1, \ldots, n.$$

Scalar multiplication. Consider an arbitrary scalar, c, and the $p \times n$ matrix $\mathbf{A}$. Scalar multiplication by a matrix then gives a new matrix in which each element of the matrix $\mathbf{A}$ is multiplied individually by the scalar c. Specifically, $c\mathbf{A} = \mathbf{A}c = \mathbf{G}$, where each element of $\mathbf{G} = \{g_{k\ell}\}$ is given by $g_{k\ell} = ca_{k\ell}$, for $k = 1, \ldots, p$ and $\ell = 1, \ldots, n$.

Matrix subtraction. As with matrix addition, matrix subtraction is defined for two matrices that have the same dimension. Consider the two $p \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$, with elements $\{a_{k\ell}\}$ and $\{b_{k\ell}\}$, for $k = 1, \ldots, p$ and $\ell = 1, \ldots, n$, respectively. The matrix difference between $\mathbf{A}$ and $\mathbf{B}$ is then given by

$$\mathbf{C} = \{c_{k\ell}\} = \mathbf{A} - \mathbf{B} = \mathbf{A} + (-1)\mathbf{B},$$

where it can be seen that the elements of $\mathbf{C}$ are given by $c_{k\ell} = a_{k\ell} - b_{k\ell}$, for $k = 1, \ldots, p$ and $\ell = 1, \ldots, n$. Thus, matrix subtraction is just a combination of matrix addition and scalar multiplication (by -1).

Matrix multiplication. The product of the $p \times n$ matrix $\mathbf{A}$ and $n \times m$ matrix $\mathbf{B}$ is given by the $p \times m$ matrix $\mathbf{C}$, where $\mathbf{C} = \{c_{kj}\} = \mathbf{AB}$, with

$$c_{kj} = \sum_{\ell=1}^{n} a_{k\ell} b_{\ell j}, \quad k = 1, \ldots, p;\ j = 1, \ldots, m.$$

Thus, for the matrix product $\mathbf{AB}$ to exist, the number of columns in $\mathbf{A}$ must equal the number of rows in $\mathbf{B}$; so $\mathbf{C}$ always has the number of rows that are in $\mathbf{A}$ and the number of columns that are in $\mathbf{B}$.

Orthogonal matrix. A square $p \times p$ matrix $\mathbf{A}$ is said to be *orthogonal* if $\mathbf{AA}' = \mathbf{A}'\mathbf{A} = \mathbf{I}_p$.

Vector inner product. As a special case of matrix multiplication, consider two vectors, $\mathbf{a}$ and $\mathbf{b}$, both of length p. The *inner product* of $\mathbf{a}$ and $\mathbf{b}$ is given by the scalar $\mathbf{a}'\mathbf{b} = \mathbf{b}'\mathbf{a} \equiv \sum_{k=1}^{p} a_k b_k$.

Vector outer product. For another special case of matrix multiplication, consider a p-dimensional vector $\mathbf{a}$ and a q-dimensional vector $\mathbf{b}$. The *outer product* $\mathbf{ab}'$ is given by the $p \times q$ matrix

$$\mathbf{ab}' \equiv \begin{bmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_q \\ a_2b_1 & a_2b_2 & \cdots & a_2b_q \\ \vdots & \vdots & & \vdots \\ a_pb_1 & a_pb_2 & \cdots & a_pb_q \end{bmatrix}.$$

Note that (in general) $\mathbf{ab}' \neq \mathbf{b}'\mathbf{a}$.

Kronecker product. Consider two matrices, an $n_a \times m_a$ matrix, $\mathbf{A}$, and an $n_b \times m_b$ matrix, $\mathbf{B}$. The Kronecker product of $\mathbf{A}$ and $\mathbf{B}$ is given by the $n_an_b \times m_am_b$ matrix $\mathbf{A} \otimes \mathbf{B}$ defined as

$$\mathbf{A} \otimes \mathbf{B} \equiv \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1m_a}\mathbf{B} \\ \vdots & \vdots & \vdots \\ a_{n_a1}\mathbf{B} & \cdots & a_{n_am_a}\mathbf{B} \end{bmatrix}.$$

If $\mathbf{A}$ is $n_a \times n_a$ and $\mathbf{B}$ is $n_b \times n_b$, the inverse and determinant of the Kronecker product can be expressed in terms of the Kronecker product of the inverses and determinants of the individual matrices, respectively:

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1},$$

$$|\mathbf{A} \otimes \mathbf{B}| = |\mathbf{A}|^{n_b}\,|\mathbf{B}|^{n_a}.$$

Euclidean norm. Consider the p-dimensional real-valued vector $\mathbf{a} = [a_1, a_2, \ldots, a_p]'$. The Euclidean norm is simply the Euclidean distance in p-dimensional space, given by

$$||\mathbf{a}|| \equiv \sqrt{\mathbf{a}'\mathbf{a}} \equiv \sqrt{\sum_{k=1}^{p} a_k^2}.$$

Symmetric matrix. A matrix $\mathbf{A}$ is said to be *symmetric* if $\mathbf{A}' = \mathbf{A}$.

Diagonal matrix. Consider the $p \times p$ matrix $\mathbf{A}$. The (main) diagonal elements of this matrix are given by the vector $[a_{11}, a_{22}, \ldots, a_{pp}]'$. Sometimes it is helpful to use a shorthand notation to construct a matrix with specific elements of a vector on the main diagonal and zeros for all other elements. For example,

$$\mathrm{diag}(b_1, b_2, \ldots, b_q) \equiv \begin{bmatrix} b_1 & 0 & 0 & \cdots & 0 \\ 0 & b_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & b_q \end{bmatrix}.$$

Trace of a matrix. Let $\mathbf{A}$ be a $p \times p$ square matrix. We then define the *trace* of this matrix, denoted $\mathrm{trace}(\mathbf{A})$ (or $\mathrm{tr}(\mathbf{A})$) as the sum of the diagonal elements of $\mathbf{A}$; that is,

$$\mathrm{trace}(\mathbf{A}) = \sum_{k=1}^{p} a_{kk}.$$

Non-negative-definite and positive-definite matrices. Consider a $p \times p$ symmetric and real-valued matrix, $\mathbf{A}$. If, for any non-zero real-valued vector $\mathbf{x}$, the scalar given by the *quadratic form* $\mathbf{x}'\mathbf{A}\mathbf{x}$ is non-negative, we say $\mathbf{A}$ is a *non-negative-definite* matrix. Similarly, if $\mathbf{x}'\mathbf{A}\mathbf{x}$ is strictly positive for any $\mathbf{x} \neq \mathbf{0}$, we say that $\mathbf{A}$ is a *positive-definite* matrix.

Matrix inverse. Consider the $p \times p$ square matrix, $\mathbf{A}$. If it exists, the matrix $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{BA} = \mathbf{I}_p$ is known as the *inverse matrix* of $\mathbf{A}$, and it is denoted by $\mathbf{A}^{-1}$. Thus, $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_p$. If the inverse exists, we say that the matrix is *invertible*. Not every square matrix has an inverse, but every positive-definite matrix is invertible (and, the inverse matrix is also positive-definite).

Matrix square root. Let $\mathbf{A}$ be a $p \times p$ positive-definite matrix. Then there exists a matrix $\mathbf{B}$ such that $\mathbf{A} = \mathbf{BB} \equiv \mathbf{B}^2$ and we say that $\mathbf{B}$ is the *matrix square root* of $\mathbf{A}$ and denote it by $\mathbf{A}^{1/2}$. The matrix square root of a positive-definite matrix is also positive-definite and we can write the inverse matrix as $\mathbf{A}^{-1} = \mathbf{A}^{-1/2}\mathbf{A}^{-1/2}$, where $\mathbf{A}^{-1/2}$ is the inverse of $\mathbf{A}^{1/2}$.

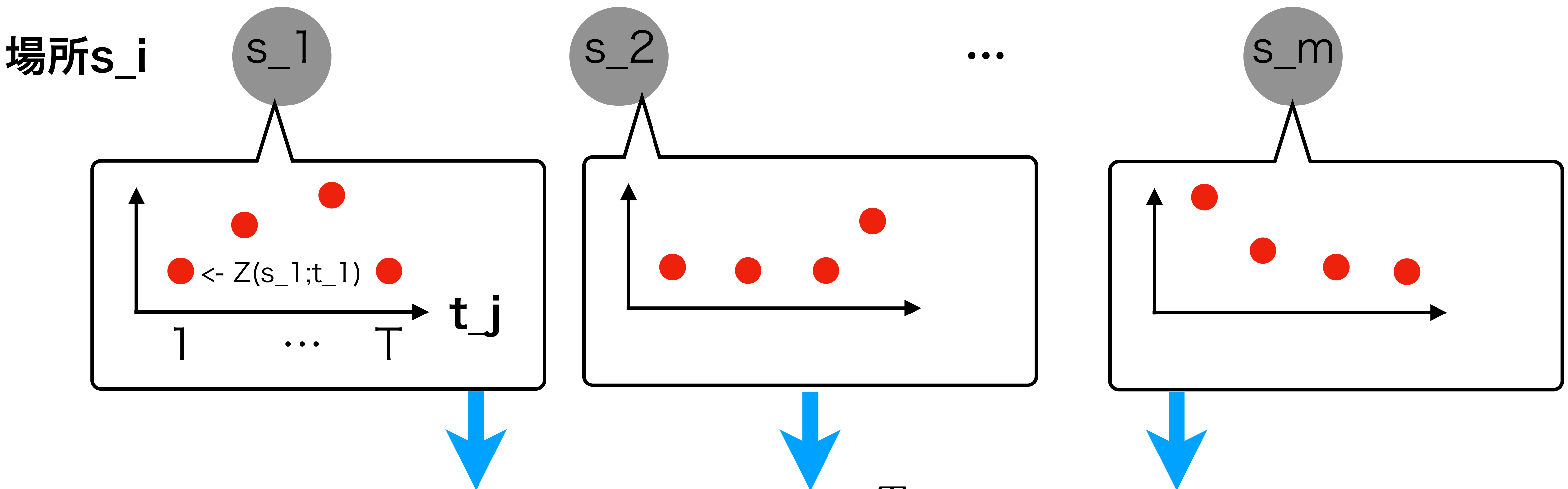
Spectral decomposition. Let $\mathbf{A}$ be a $p \times p$ symmetric matrix of real values. This matrix can be decomposed as

$$\mathbf{A} = \sum_{k=1}^{p} \lambda_k \boldsymbol{\phi}_k \boldsymbol{\phi}_k' = \boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{\Phi}',$$

where $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$, $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_p]$, and $\{\lambda_k\}$ are called the *eigenvalues* that are associated with the *eigenvectors*, $\{\boldsymbol{\phi}_k\}$, $k = 1, \ldots, p$, which are orthogonal (i.e., $\boldsymbol{\Phi}\boldsymbol{\Phi}' = \boldsymbol{\Phi}'\boldsymbol{\Phi} = \mathbf{I}_p$). Note that for a symmetric non-negative-definite matrix $\mathbf{A}$, $\lambda_k \geq 0$, and for a symmetric positive-definite matrix $\mathbf{A}$, $\lambda_k > 0$, for all $k = 1, \ldots, p$. The matrix square root and its inverse can be written as $\mathbf{A}^{1/2} = \boldsymbol{\Phi}\,\mathrm{diag}(\lambda_1^{1/2}, \ldots, \lambda_p^{1/2})\boldsymbol{\Phi}'$ and $\mathbf{A}^{-1/2} = \boldsymbol{\Phi}\,\mathrm{diag}(\lambda_1^{-1/2}, \ldots, \lambda_p^{-1/2})\boldsymbol{\Phi}'$, respectively.

Singular value decomposition (SVD). Let $\mathbf{A}$ be a $p \times n$ matrix of real values. Then the matrix $\mathbf{A}$ can be decomposed as $\mathbf{A} = \mathbf{UDV}'$, where $\mathbf{U}$ and $\mathbf{V}$ are $p \times p$ and $n \times n$ orthogonal matrices, respectively. In addition, the $p \times n$ matrix $\mathbf{D}$ contains all zeros except for the (k, k)th non-negative elements, $\{d_k : \ k = 1, 2, \ldots, \min(p, n)\}$, which are known as *singular values*.

2.4.1 Empirical spatial means and covariances (p.33-)

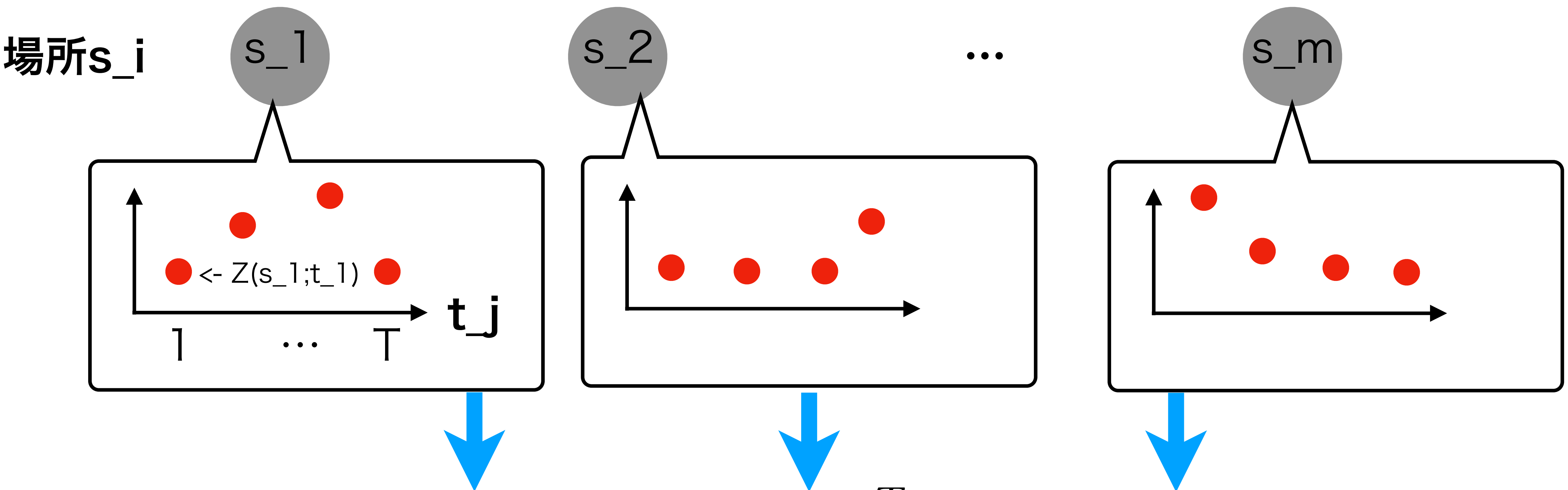


Spatial mean

$$\widehat{\mu}_{z,s}(\mathbf{s}_i) \equiv \frac{1}{T}\sum_{j=1}^{T} Z(\mathbf{s}_i;t_j)$$

$$\widehat{\boldsymbol{\mu}}_{z,s} \equiv \begin{bmatrix} \widehat{\mu}_{z,s}(\mathbf{s}_1) \\ \vdots \\ \widehat{\mu}_{z,s}(\mathbf{s}_m) \end{bmatrix} = \begin{bmatrix} \frac{1}{T}\sum_{j=1}^{T} Z(\mathbf{s}_1;t_j) \\ \vdots \\ \frac{1}{T}\sum_{j=1}^{T} Z(\mathbf{s}_m;t_j) \end{bmatrix} = \frac{1}{T}\sum_{j=1}^{T} \mathbf{Z}_{t_j}, \qquad (2.1)$$

2.4.1 Empirical spatial means and covariances (p.33-)



Spatial mean

$$\widehat{\mu}_{z,s}(\mathbf{s}_i) \equiv \frac{1}{T}\sum_{j=1}^{T} Z(\mathbf{s}_i; t_j)$$

Time mean

$$\widehat{\mu}_{z,t}(t_j) \equiv \frac{1}{m}\sum_{i=1}^{m} Z(\mathbf{s}_i; t_j). \qquad (2.2)$$

t=1のときの(空間)平均
t=2 〃
⋮
t=T 〃
っていうベクトル

Spatial covariability, cross- covariance matrix

Spatial covariability

$$\widehat{C}_z^{(\tau)}(\mathbf{s}_i, \mathbf{s}_k) \equiv \frac{1}{T-\tau} \sum_{j=\tau+1}^{T} (Z(\mathbf{s}_i; t_j) - \widehat{\mu}_{z,s}(\mathbf{s}_i))(Z(\mathbf{s}_k; t_j - \tau) - \widehat{\mu}_{z,s}(\mathbf{s}_k)), \qquad (2.3)$$

$$\Leftrightarrow \; \widehat{\mathbf{C}}_z^{(\tau)} \equiv \frac{1}{T-\tau} \sum_{j=\tau+1}^{T} (\mathbf{Z}_{t_j} - \widehat{\boldsymbol{\mu}}_{z,s})(\mathbf{Z}_{t_j-\tau} - \widehat{\boldsymbol{\mu}}_{z,s})'; \quad \tau = 0, 1, \ldots, T-1. \qquad (2.4)$$

lag 0

lag 1

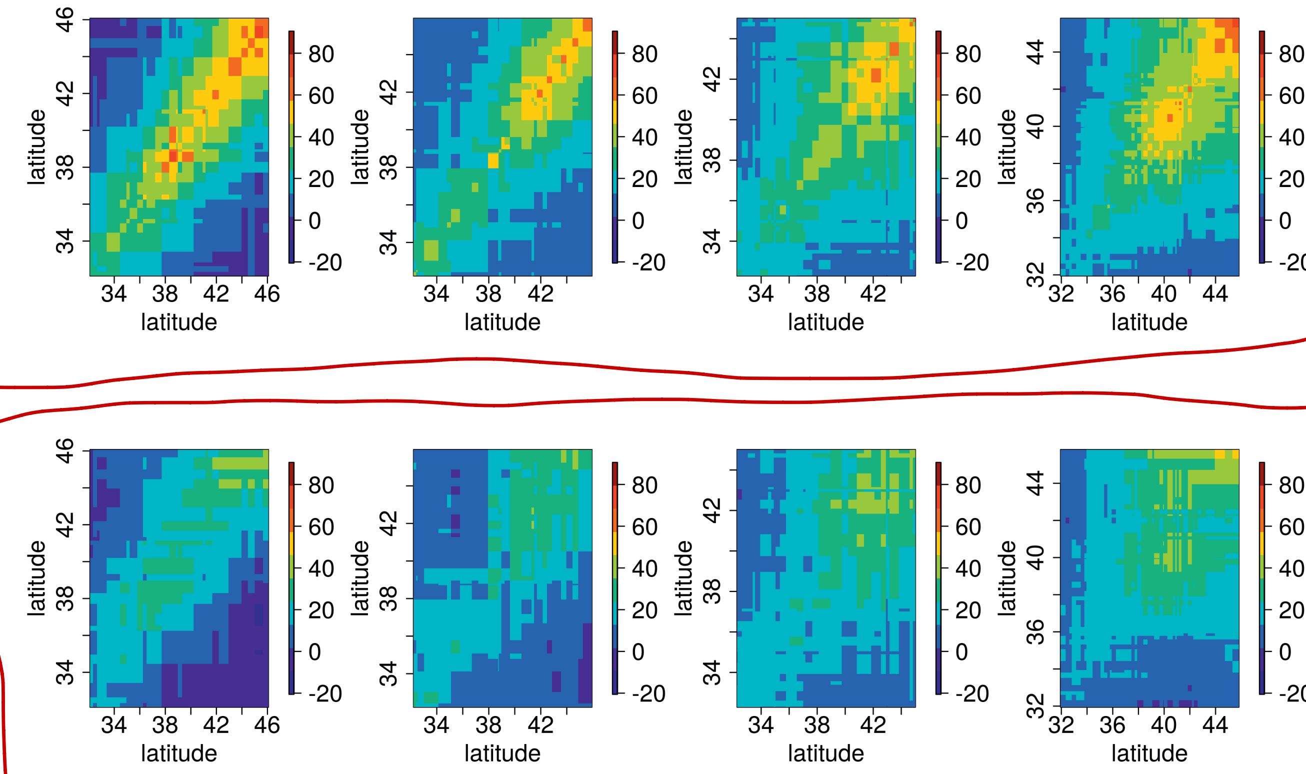


Figure 2.16: Maximum temperature lag-0 (top) and lag-1 (bottom) empirical spatial covariance plots for four longitudinal strips (from left to right, $[-100,-95), [-95,-90), [-90,-85), [-85,-80)$ degrees) in which the domain of interest is subdivided.

Spatial covariability

$$\widehat{C}_z^{(\tau)}(\mathbf{s}_i,\mathbf{s}_k) \equiv \frac{1}{T-\tau}\sum_{j=\tau+1}^{T}(Z(\mathbf{s}_i;t_j)-\widehat{\mu}_{z,s}(\mathbf{s}_i))(Z(\mathbf{s}_k;t_j-\tau)-\widehat{\mu}_{z,s}(\mathbf{s}_k)), \qquad (2.3)$$

$$\Leftrightarrow\ \widehat{\mathbf{C}}_z^{(\tau)} \equiv \frac{1}{T-\tau}\sum_{j=\tau+1}^{T}(\mathbf{Z}_{t_j}-\widehat{\boldsymbol{\mu}}_{z,s})(\mathbf{Z}_{t_j-\tau}-\widehat{\boldsymbol{\mu}}_{z,s})'; \quad \tau = 0,1,\ldots,T-1. \qquad (2.4)$$

Spatial covariability, cross- covariance matrix

Spatial covariability

$$\widehat{C}_z^{(\tau)}(\mathbf{s}_i, \mathbf{s}_k) \equiv \frac{1}{T-\tau} \sum_{j=\tau+1}^{T} (Z(\mathbf{s}_i; t_j) - \widehat{\mu}_{z,s}(\mathbf{s}_i))(Z(\mathbf{s}_k; t_j - \tau) - \widehat{\mu}_{z,s}(\mathbf{s}_k)), \qquad (2.3)$$

$$\Leftrightarrow \ \widehat{\mathbf{C}}_z^{(\tau)} \equiv \frac{1}{T-\tau} \sum_{j=\tau+1}^{T} (\mathbf{Z}_{t_j} - \widehat{\boldsymbol{\mu}}_{z,s})(\mathbf{Z}_{t_j-\tau} - \widehat{\boldsymbol{\mu}}_{z,s})'; \quad \tau = 0, 1, \ldots, T-1. \qquad (2.4)$$

Cross-covariance matrix

$$\widehat{\mathbf{C}}_{z,x}^{(\tau)} \equiv \frac{1}{T-\tau} \sum_{j=\tau+1}^{T} (\mathbf{Z}_{t_j} - \widehat{\boldsymbol{\mu}}_{z,s})(\mathbf{X}_{t_j-\tau} - \widehat{\boldsymbol{\mu}}_{x,s})', \qquad (2.5)$$

2.4.2 Spatio-temporal covariograms and semivariograms (p.36-)

Spatio-temporal covariance function

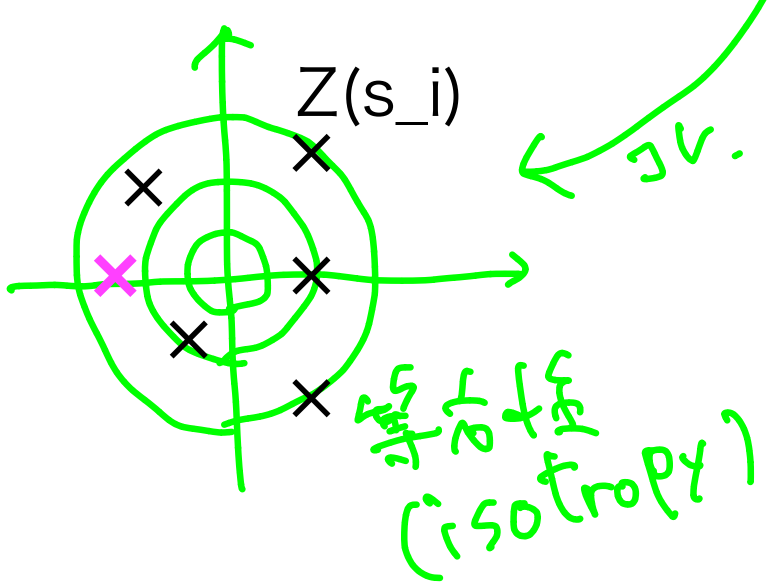
$$\widehat{C}_z(\mathbf{h};\tau) = \frac{1}{|N_{\mathbf{s}}(\mathbf{h})|}\frac{1}{|N_t(\tau)|}\sum_{\mathbf{s}_i,\mathbf{s}_k \in N_{\mathbf{s}}(\mathbf{h})}\sum_{t_j,t_\ell \in N_t(\tau)} (Z(\mathbf{s}_i;t_j) - \widehat{\mu}_{z,s}(\mathbf{s}_i))(Z(\mathbf{s}_k;t_\ell) - \widehat{\mu}_{z,s}(\mathbf{s}_k)), \qquad (2.6)$$

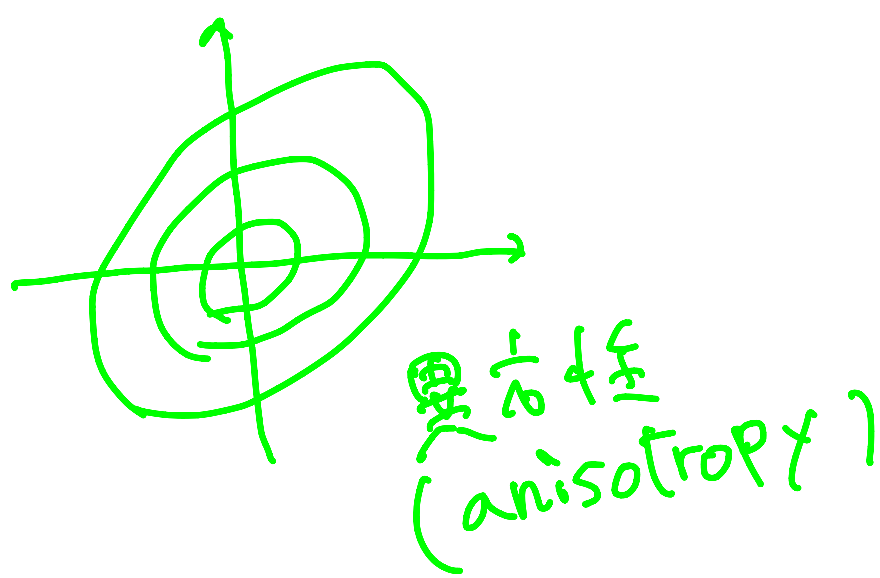
ラグが h or τ
の時のペアの数

s_iの系列
⋮
s_mの系列

Under *isotropy*, one often considers the lag only as a function of distance, $h = ||\mathbf{h}||$, where $||\cdot||$ is the Euclidean norm (see Appendix A).

コレ.





×の値を予測するには‥
『時空間的な距離にともなって観測値がどう変わるのか？（つまり時空間自己相関は？）』
-> セミバリオグラム

$$\underbrace{\frac{1}{2}\mathrm{var}\{Z(s+h;t+\tau)\}}_{=C(0;0)\ \text{一定値}} + \underbrace{\frac{1}{2}\mathrm{var}\{Z(s;t)\}}_{=C(0;0)} - \underbrace{\mathrm{cov}[Z(s+h;t+\tau), Z(s;t)]}_{C((s+h;t+\tau)-(s;t)) = C(h;\tau)}$$

Zの共分散は相対的な位置h,τにのみ依存するから（2次定常性）

Technical Note 2.1: Semivariogram

The semivariogram is defined as

セミ

$$\gamma_z(\mathbf{s}_i, \mathbf{s}_k; t_j, t_\ell) \equiv \frac{1}{2}\mathrm{var}(Z(\mathbf{s}_i; t_j) - Z(\mathbf{s}_k; t_\ell)).$$

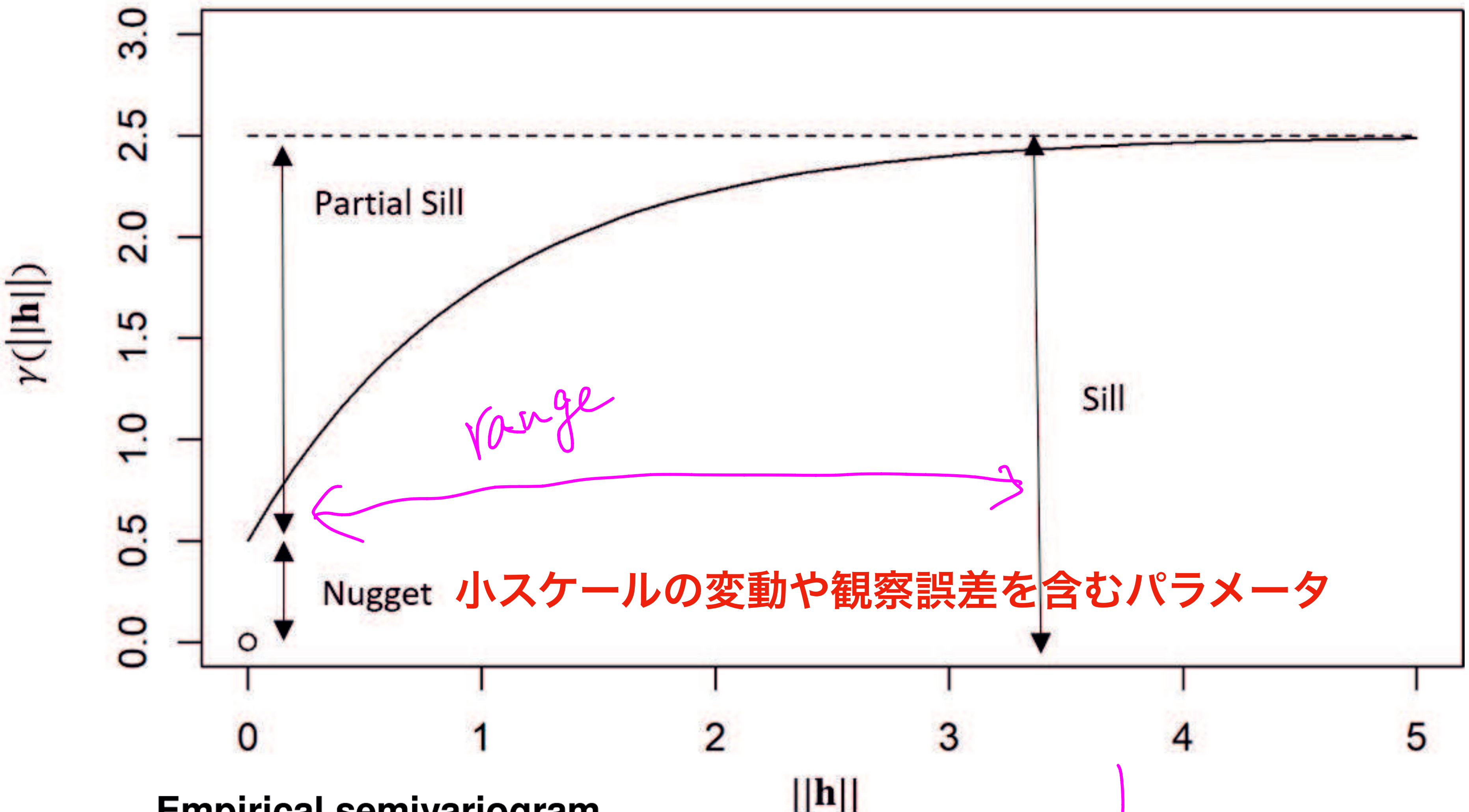
In the case where the covariance depends only on displacements in space and differences in time, this can be written as

$$\begin{aligned} \gamma_z(\mathbf{h};\tau) &= \frac{1}{2}\mathrm{var}(Z(\mathbf{s}+\mathbf{h};t+\tau) - Z(\mathbf{s};t)) \\ &= C_z(\mathbf{0};0) - \mathrm{cov}(Z(\mathbf{s}+\mathbf{h};t+\tau), Z(\mathbf{s};t)) \\ &= C_z(\mathbf{0};0) - C_z(\mathbf{h};\tau), \end{aligned} \qquad (2.7)$$

where $\mathbf{h} = \mathbf{s}_k - \mathbf{s}_i$ is a spatial lag and $\tau = t_\ell - t_j$ is a temporal lag.

Now, (2.7) does not always hold. It is possible that γ_z is a function of spatial lag $\mathbf{h}$ and temporal lag τ, but there is no stationary covariance function $C_z(\mathbf{h};\tau)$. We generally

or コバリオグラム



時空間のセミバリオグラム

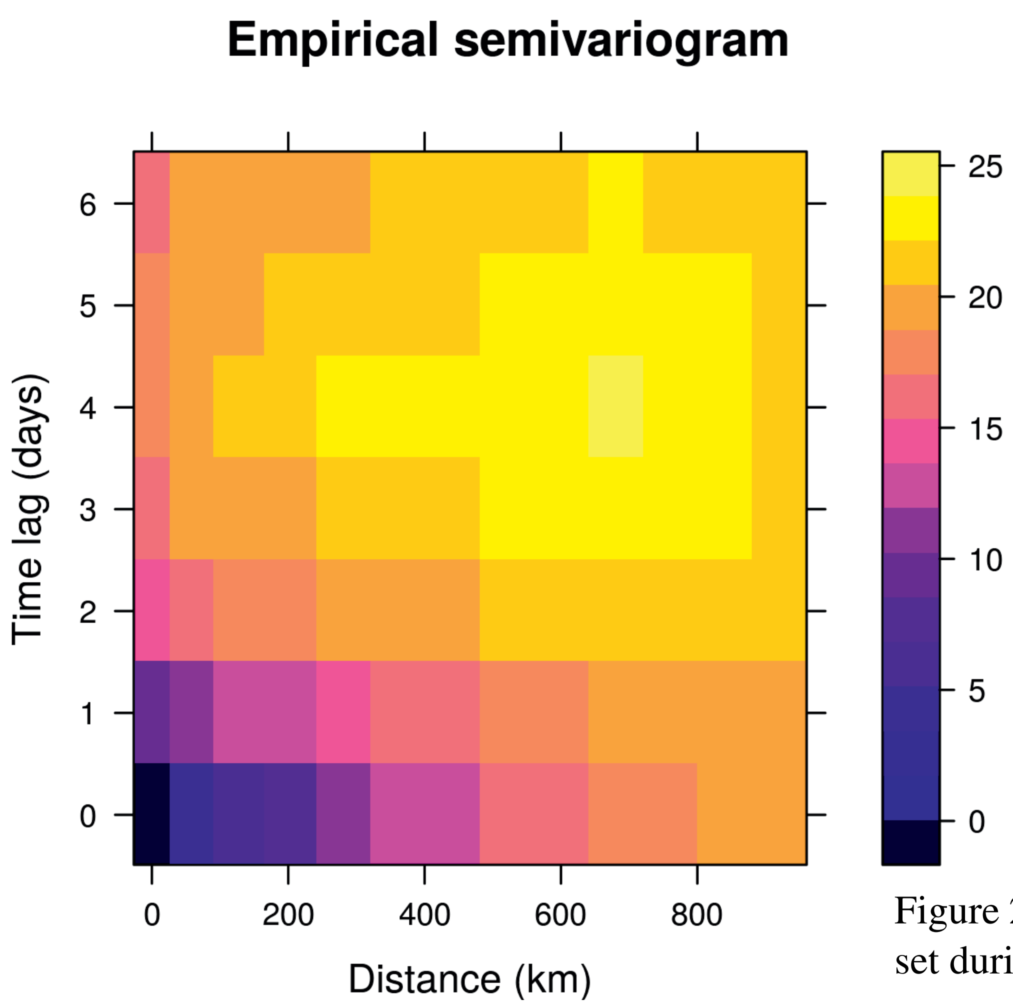


Figure 2.17: Empirical spatio-temporal semivariogram of daily `Tmax` from the NOAA data set during July 2003, computed using the function **`variogram`** in **gstat**.

2.4.3 Empirical orthogonal functions (p.37-)

- PCA

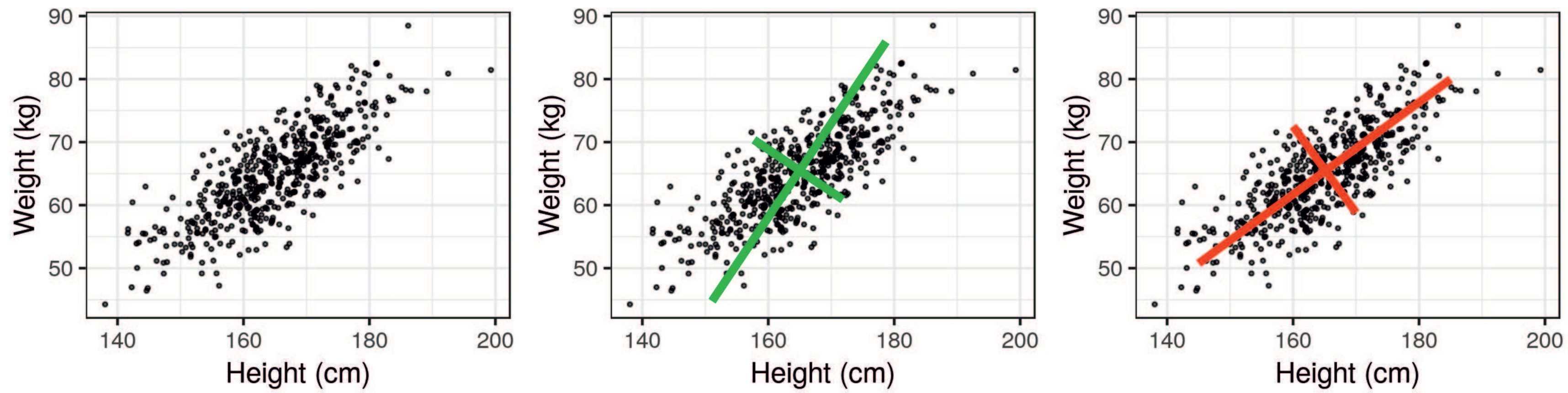


Figure 2.18: Simulated height (in cm) versus weight (in kg) for $m = 500$ females in the USA (left) with two orthogonal projections (center and right). The right panel shows the optimal PCA projection.

$$a_1 = w_{11}x_1 + w_{12}x_2$$
$$a_2 = w_{21}x_1 + w_{22}x_2$$

As an example, consider the variance–covariance matrix associated with the simulated height and weight traits, where $p = 2$:

$$\widehat{\mathbf{C}}_z^{(\tau)} \equiv \frac{1}{T-\tau} \sum_{j=\tau+1}^{T} (\mathbf{Z}_{t_j} - \widehat{\boldsymbol{\mu}}_{z,s})(\mathbf{Z}_{t_j-\tau} - \widehat{\boldsymbol{\mu}}_{z,s})'; \quad \tau = 0, 1, \ldots, T-1. \qquad (2.4)$$

$\tau = 0$

$$\mathbf{C}_x = \begin{pmatrix} 81 & 50 \\ 50 & 49 \end{pmatrix}.$$

Then $\mathbf{W}$ and $\boldsymbol{\Lambda}_x$ are given (using the function **eigen** in R) by

w_{11} w_{21}

$$\mathbf{W} = \begin{pmatrix} -0.8077 & 0.5896 \\ -0.5896 & -0.8077 \end{pmatrix}, \quad \boldsymbol{\Lambda}_x = \begin{pmatrix} 117.5 & 0 \\ 0 & 12.5 \end{pmatrix}.$$

w_{12} w_{22}

Let $\mathbf{Z}_{t_j} \equiv (Z(\mathbf{s}_1;t_j),\ldots,Z(\mathbf{s}_m;t_j))'$ for $j = 1,\ldots,T$. Using (2.4) to estimate the lag-0 spatial covariance matrix, $\widehat{\mathbf{C}}_z^{(0)}$ (which is symmetric and non-negative-definite), the PCA decomposition is given by the spectral decomposition

"EOFs" Figs. 2.20a, 2.21a

$$\widehat{\mathbf{C}}_z^{(0)} = \mathbf{\Psi}\mathbf{\Lambda}\mathbf{\Psi}', \qquad \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix} \quad \lambda \geq 0 \tag{2.9}$$

where $\mathbf{\Psi} \equiv (\boldsymbol{\psi}_1,\ldots,\boldsymbol{\psi}_m)$ is a matrix of spatially indexed eigenvectors given by the vectors $\boxed{\boldsymbol{\psi}_k \equiv (\psi_k(\mathbf{s}_1),\ldots,\psi_k(\mathbf{s}_m))'}$ for $k = 1,\ldots,m$, and $\mathbf{\Lambda} \equiv \mathrm{diag}(\lambda_1,\ldots,\lambda_m)$ is a diagonal matrix of corresponding non-negative eigenvalues (decreasing down the diagonal). The

***kth* PC time series**

$$a_k(t_j) \equiv \boldsymbol{\psi}_k' \mathbf{Z}_{t_j}$$

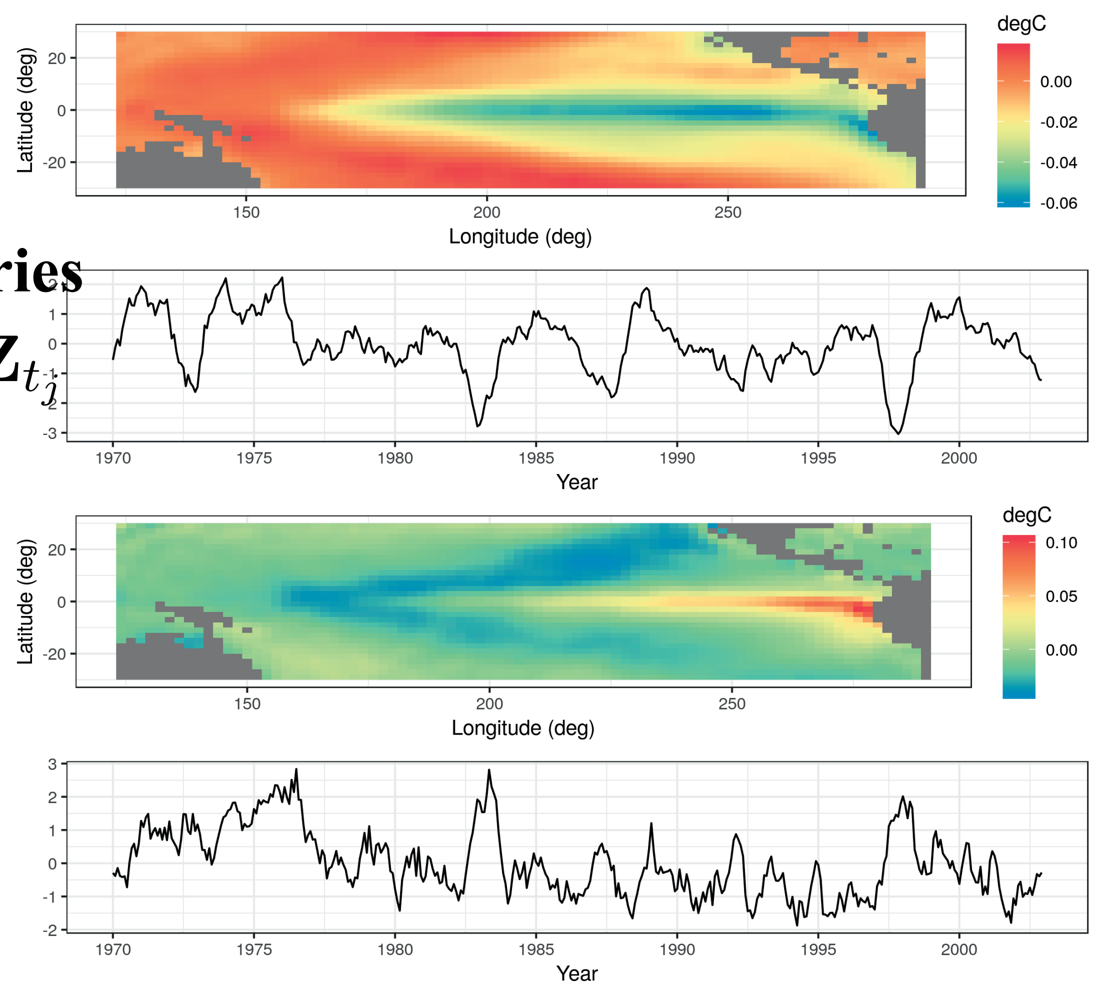


Figure 2.20: The first two empirical orthogonal functions and normalized principal-component time series for the SST data set obtained using an SVD of a space-wide matrix.

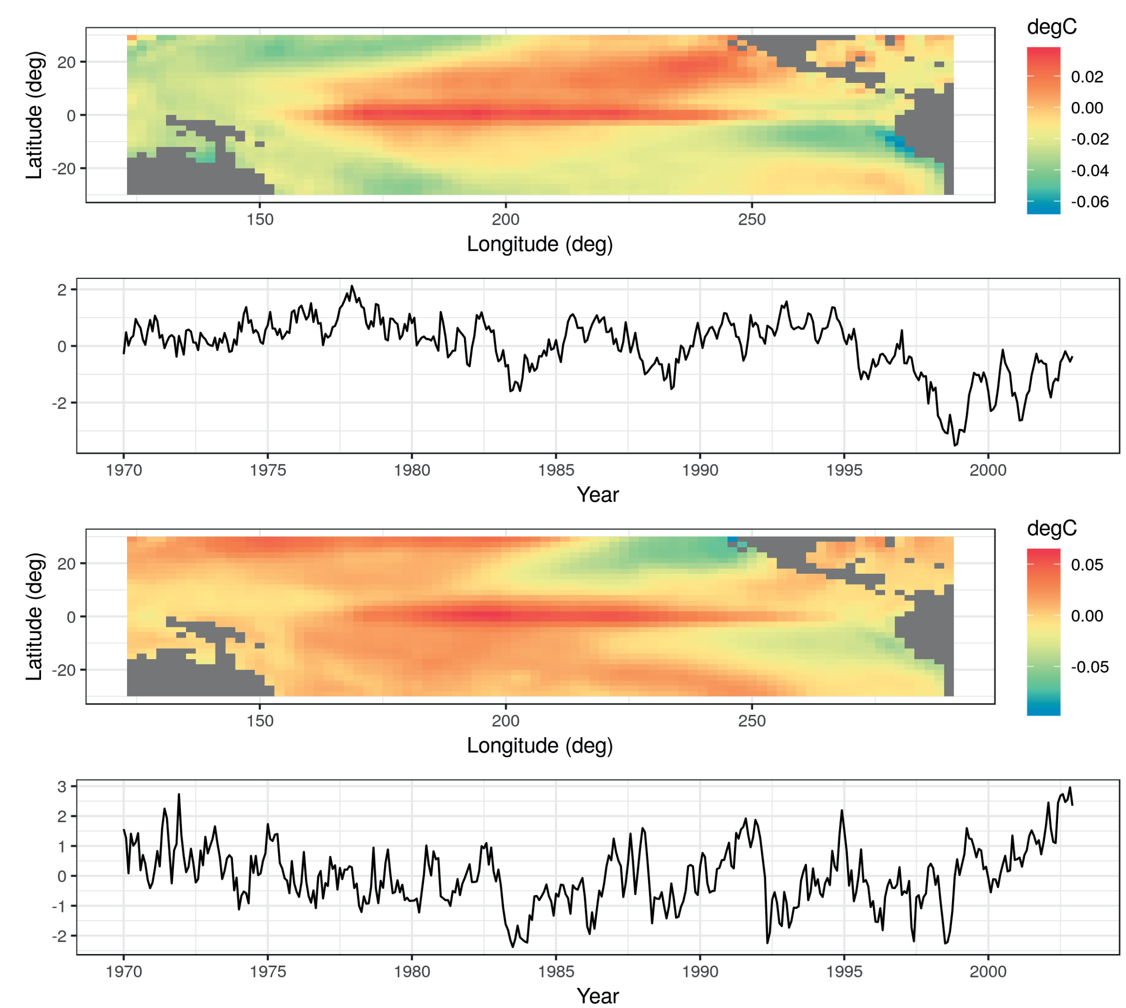


Figure 2.21: The third and fourth empirical orthogonal functions and normalized principal-component time series for the SST data set obtained using an SVD of a space-wide matrix.

2.4.3 Spatio-temporal canonical correlation analysis (p.47-)

Assume that we have two data sets that have the same temporal domain of interest but potentially different spatial domains. In particular, consider the data sets given by the collection of spatial vectors $\{\mathbf{Z}_{t_j} \equiv (Z(\mathbf{s}_1; t_j), \ldots, Z(\mathbf{s}_m; t_j))' : j = 1, \ldots, T\}$, and $\{\mathbf{X}_{t_j} \equiv (X(\mathbf{r}_1; t_j), \ldots, X(\mathbf{r}_n; t_j))' : j = 1, \ldots, T\}$. Now, consider the two new variables

that are linear combinations of $\mathbf{Z}_{t_j}$ and $\mathbf{X}_{t_j}$, respectively:

$$a_k(t_j) = \sum_{i=1}^{m} \xi_{ik}\, Z(\mathbf{s}_i; t_j) = \boldsymbol{\xi}_k' \mathbf{Z}_{t_j}, \tag{2.13}$$

$$b_k(t_j) = \sum_{\ell=1}^{n} \psi_{\ell k}\, X(\mathbf{r}_\ell; t_j) = \boldsymbol{\psi}_k' \mathbf{X}_{t_j}. \tag{2.14}$$

For suitable choices of weights (see below), the kth *canonical correlation*, for $k = 1, 2, \ldots, \min\{n, m\}$, is then simply the correlation between a_k and b_k,

$$r_k \equiv \mathrm{corr}(a_k, b_k) = \frac{\mathrm{cov}(a_k, b_k)}{\sqrt{\mathrm{var}(a_k)}\sqrt{\mathrm{var}(b_k)}},$$

which can also be written as

$$r_k = \frac{\boldsymbol{\xi}_k' \mathbf{C}_{z,x}^{(0)} \boldsymbol{\psi}_k}{(\boldsymbol{\xi}_k' \mathbf{C}_z^{(0)} \boldsymbol{\xi}_k)^{1/2} (\boldsymbol{\psi}_k' \mathbf{C}_x^{(0)} \boldsymbol{\psi}_k)^{1/2}}, \tag{2.15}$$

2.4.3 Spatio-temporal canonical correlation analysis (p.47-)

As an example of ST-CCA, we consider a one-field ST-CCA on the SST data set. In particular, we are interested in forecasting SST seven months in the future, so we let the data $\mathbf{X}$ be the lag $\tau = 7$ month SST data and the data $\mathbf{Z}$ be the same SSTs with no lag.

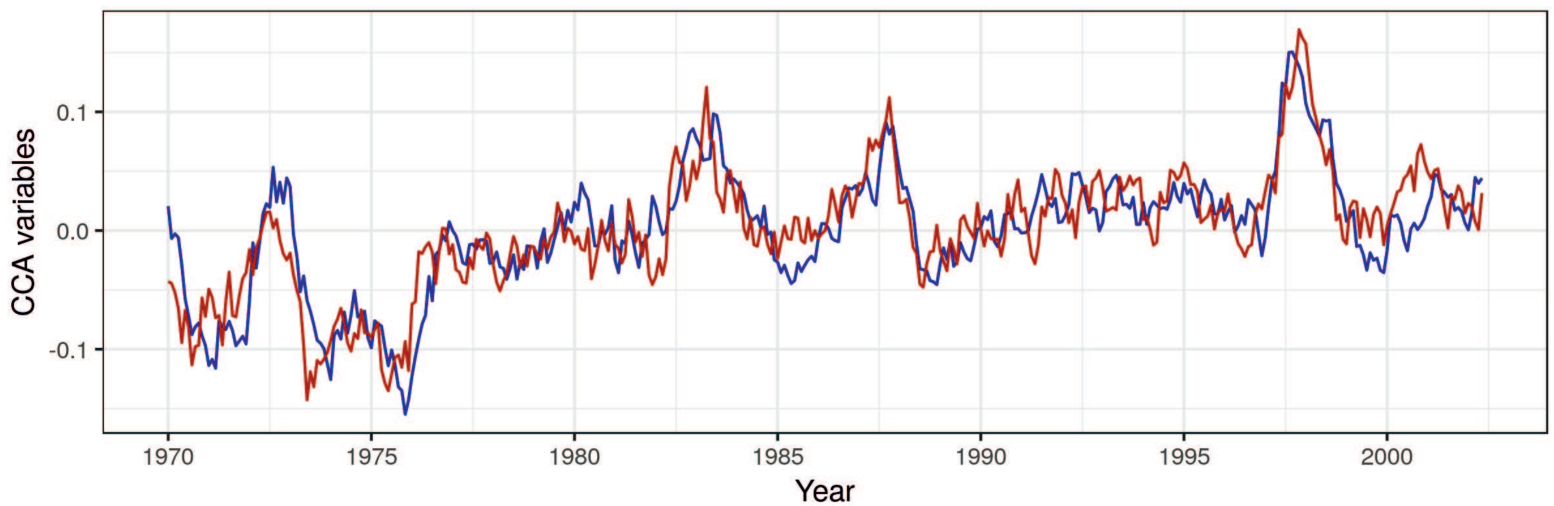


Figure 2.23: Time series of the first canonical variables, $\{a_1, b_1\}$, for $\tau = 7$ month lagged monthly SST anomalies at time $t_j - \tau$ (blue) and those at time t_j (red).

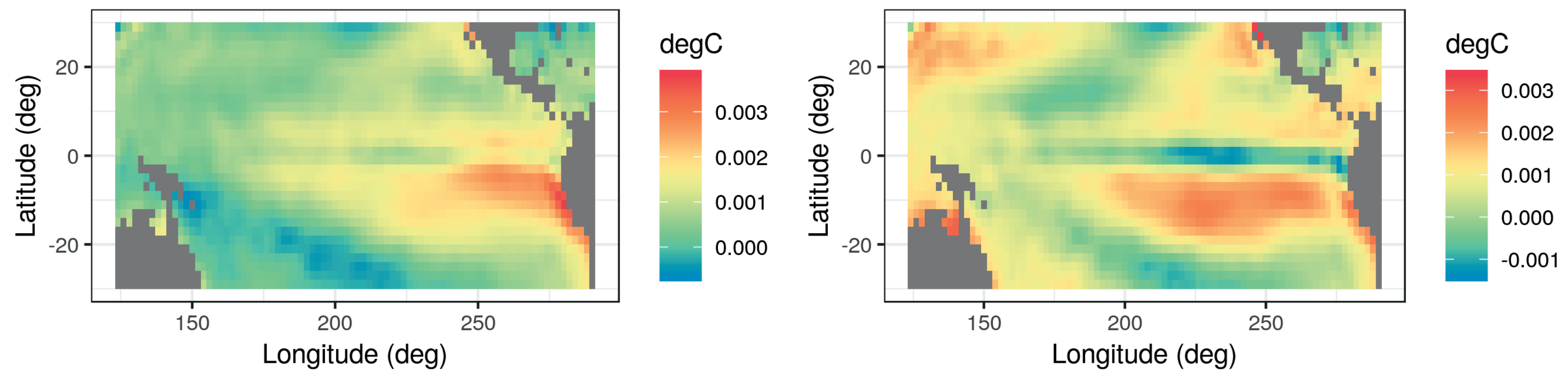


Figure 2.24: Spatial-weights maps corresponding to the linear combination of EOFs used to construct the canonical variables for SST data lagged $\tau = 7$ months (left) and the unlagged SST data (right).

2.5 Chapter 2 wrap-up

1. Rでデータ構造を見る
2. 絵を描く
3. 種々の方法でデータを要約する

作図してみよー: 準備

```
#packages
require(STRbook)
require(tidyr)
require(dplyr)
require(plyr)
require(ggplot2)
require(CCA)
require(gstat)
require(sp)
require(spacetime)
require(fields)
```

```
#####prepare for data#####
set.seed(1)

data("NOAA_df_1990", package = "STRbook")
Tmax <- filter(NOAA_df_1990,
               proc == "Tmax" &
               month %in% 5:9 &
               year == 1993)
Tmax$t <- Tmax$julian - 728049
```

Fig.2.14を描く（p.68-)

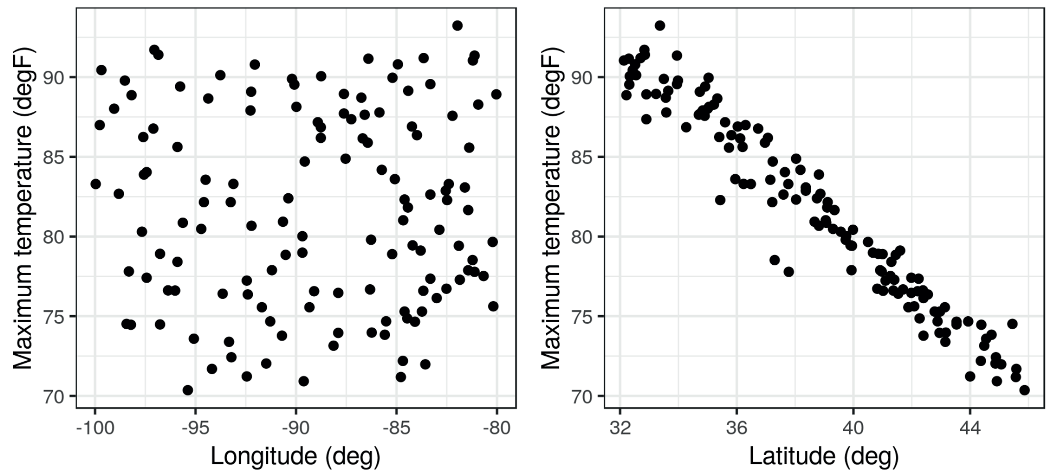


Figure 2.14: Empirical spatial mean, $\widehat{\mu}_{z,s}(\cdot)$, of Tmax (in °F) as a function of station longitude (left) and station latitude (right).

```
spat_av <- ddply(Tmax, .(lat, lon), summarize, mu_emp = mean(z))
head(spat_av)

##作図
#lat
g <- ggplot(spat_av, aes(x = lat, y = mu_emp))
p <- geom_point()
lb <- labs(x = "Latitude (deg)", y = "Maximum temperature (degF)")
lat <- g+p+lb+theme_bw()
#lon
g <- ggplot(spat_av, aes(x = lon, y = mu_emp))
p <- geom_point()
lb <- labs(x = "Latitude (deg)", y = "Maximum temperature (degF)")
lon <- g+p+lb+theme_bw()
#図をくっつける
require(gridExtra)
grid.arrange(lon, lat, ncol = 2, top = "Empirical Spatial Means")
```

Fig.2.15を描く（p.69)

```
head(Tmax)
temp_av <- ddply(Tmax, .(date), summarize, meanTmax = mean(z))
head(temp_av)
```

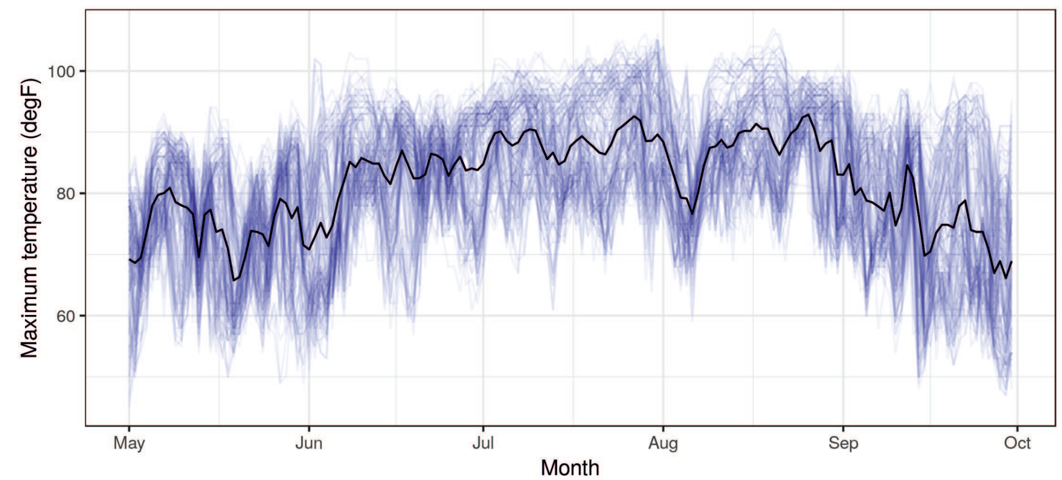


Figure 2.15: `Tmax` data (in °F), from the NOAA data set (blue lines, where each blue line corresponds to a station) and the empirical temporal mean $\widehat{\mu}_{z,t}(\cdot)$ (black line) computed from (2.2), and t is in units of days, ranging from 01 May 1993 to 30 September 1993.

```
##作図
g <- ggplot()
l1 <- geom_line(data = Tmax, aes(x = date, y = z, group = id), colour = "blue",
alpha = 0.04)
l2 <- geom_line(data = temp_av, aes(x = date, y = meanTmax))
lb <- labs(title = "Empirical Temporal Means" ,x = "Month", y = "Maximum
temperature (degF)")
g+l1+l2+lb+theme_bw()
```

Fig.2.16を描く (p.69-)

```
#トレンドの除去
Tmax$residuals <- summary(lm(z ~ lat + t + I(t^2), data = Tmax))$residuals
#必要なデータを抽出し並び替え
spat_df <- filter(Tmax, t == 1 ) %>%
           select(lon, lat) %>%
           arrange(lon, lat)
m <- nrow(spat_av)
```

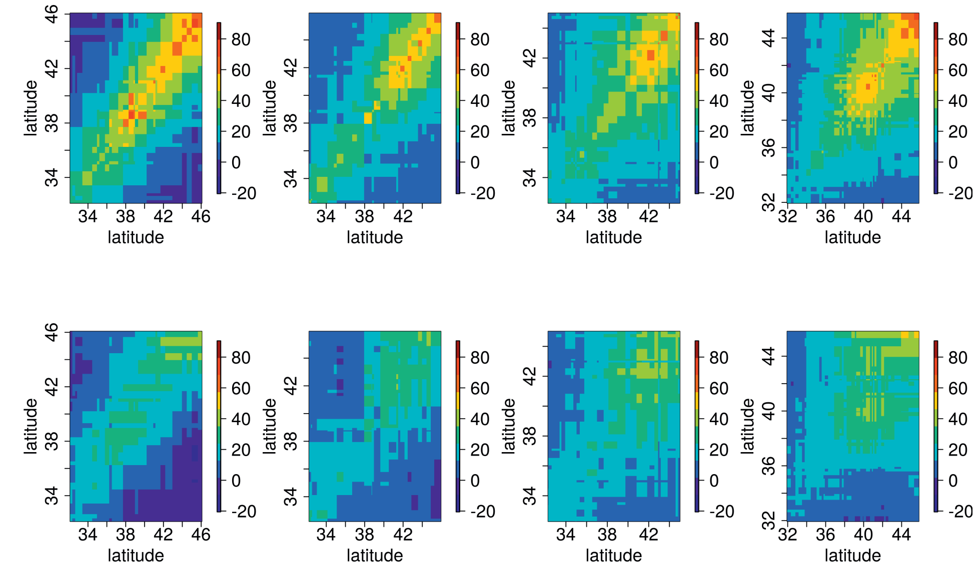


Figure 2.16: Maximum temperature lag-0 (top) and lag-1 (bottom) empirical spatial covariance plots for four longitudinal strips (from left to right, $[-100, -95), [-95, -90), [-90, -85), [-85, -80)$ degrees) in which the domain of interest is subdivided.

```
#データの形を縦長から横長に変え，行と列を入れ替える
X <- select(Tmax, lon, lat, residuals, t) %>%
     spread(key = t, value = residuals) %>% #dcastと同じ
     select(-lon, -lat) %>% #列の除去
     t() #転置

#コバリアンス行列を作成
cov0 <- cov(X, use = "complete.obs") #ラグ0
cov1 <- cov(X[-1, ], X[-nrow(X), ], use = "complete.obs") #ラグ1
```

Fig.2.16を描く (p.69-)

```
#作図のためのデータを作成
spat_df$n <- 1:nrow(spat_df)
lim_lon <- range(spat_df$lon)
lon_strips <- seq(lim_lon[1], lim_lon[2], length = 5) #-100から-80を等分割にした長さ5のベクトルを作成
spat_df$lon_strip <- cut(spat_df$lon,
                         lon_strips,
                         labels = FALSE,
                         include.lowest = TRUE) #カテゴリ分けの末端の開区間を閉区間に
summary(spat_df)#1:-99.97~-95.08, 2:-94.72~-90.08, 3:-89.98~-85.08, 4:-84.93~-80.03,
plot_cov_strips(cov0, spat_df) #図が1つしか出てこん‥
plot_cov_strips(cov1, spat_df) #図が1つしか出てこん‥
```

Fig.2.17を描く (p.71-)

```
data("STObj3", package = "STRbook")
STObj4 <- STObj3[, "1993-07-01::1993-07-31"]

require(gstat) #ダウンロードできんからセミバリオグラムが書けん
vv <- variogram(object = z ~ 1 + lat, # fixed effect component
                data = STObj4,
                width = 80,
                cutoff = 1000,
                tlags = 0.01:6.01) # 0 days to 6 days

#作図
plot(vv)
```

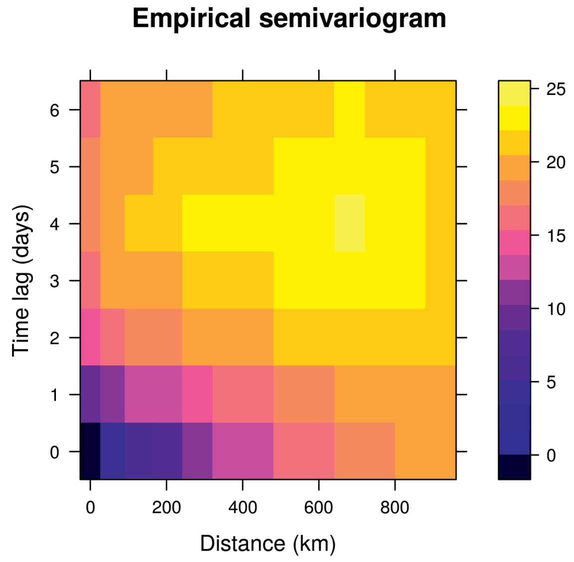


Figure 2.17: Empirical spatio-temporal semivariogram of daily `Tmax` from the NOAA data set during July 2003, computed using the function `variogram` in **gstat**.

Fig.2.20(一番上)を描く (p.72-)

```
#SSTデータ（陸地含む）
data("SSTlandmask", package = "STRbook")
data("SSTlonlat", package = "STRbook")
data("SSTdata", package = "STRbook")

delete_rows <- which(SSTlandmask == 1)
SSTdata <- SSTdata[-delete_rows, 1:396]
summary(SSTdata)

Z <- t(SSTdata)
dim(Z)

spat_mean <- apply(SSTdata, 1, mean)
nT <- ncol(SSTdata)
```

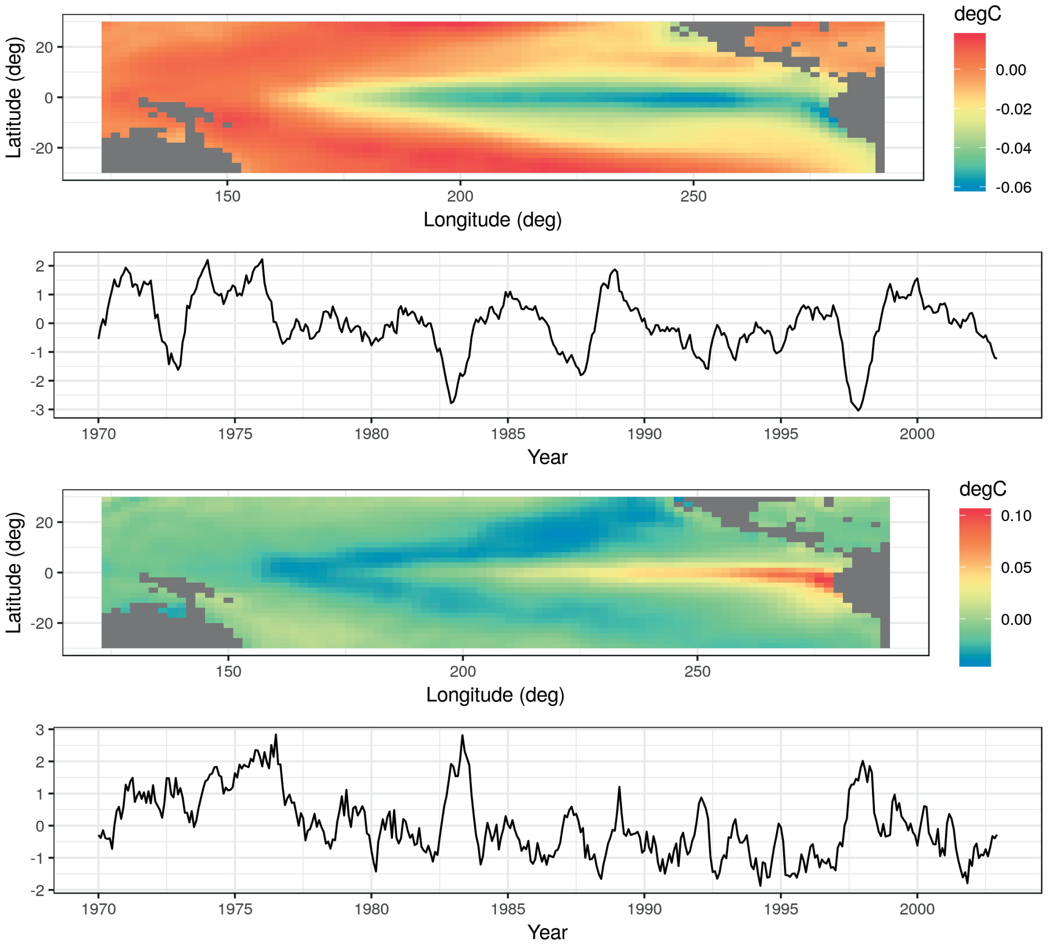


Figure 2.20: The first two empirical orthogonal functions and normalized principal-component time series for the SST data set obtained using an SVD of a space-wide matrix.

Fig.2.20(一番上)を描く (p.72-)

```
#式2.10のかっこの中を計算
Zspat_detrend <- Z - outer(rep(1, nT), spat_mean)
#式2.11を計算
Zt <- 1/sqrt(nT - 1)*Zspat_detrend
#式2.12より
E <- svd(Zt)
V <- E$v
colnames(E$v) <- paste0("EOF", 1:ncol(SSTdata))
EOFs <- cbind(SSTlonlat[-delete_rows, ], E$v)
head(EOFs[, 1:6])

#EOF1~4が各列に並んでいる（widspread）ので，EOF_kをカテゴリとして扱えるようなデータセットに変える
TS <- data.frame(E$u) %>%
      mutate(t = 1:nrow(E$u)) %>%
      gather(key = EOF, value = PC, -t) #meltやspreadと同じような感じ
#normalized
TS$nPC <- TS$PC * sqrt(nT-1)
```

space

```
## Put data into space-wide form
Z <- t(SSTdata)
dim(Z)

## [1] 396 2261
```

months

$$\widetilde{\mathbf{Z}} \equiv \frac{1}{\sqrt{T-1}}(\mathbf{Z} - \mathbf{1}_T \widehat{\boldsymbol{\mu}}'_{z,s}), \tag{2.10}$$

where $\mathbf{1}_T$ is a T-dimensional vector of ones and $\widehat{\boldsymbol{\mu}}_{z,s}$ is the spatial mean vector given by (2.1). Then it is easy to show that

$$\mathbf{C}_z^{(0)} = \widetilde{\mathbf{Z}}'\widetilde{\mathbf{Z}} = \boldsymbol{\Psi}\boldsymbol{\Lambda}\boldsymbol{\Psi}', \tag{2.11}$$

Note that Z is of size 396×2261, and it is hence in space-wide format as required. Equation (2.10) is implemented as follows.

```
## First find the matrix we need to subtract:
spat_mean <- apply(SSTdata, 1, mean)
nT <- ncol(SSTdata)

## Then subtract and standardize:
Zspat_detrend <- Z - outer(rep(1, nT), spat_mean)
Zt <- 1/sqrt(nT - 1)*Zspat_detrend
```

外積

Finally, to carry out the SVD we run

```
E <- svd(Zt)
```

The SVD returns a list E containing the matrices $\mathbf{V}$, $\mathbf{U}$, and the singular values diag($\mathbf{D}$). The matrix $\mathbf{V}$ contains the EOFs in space-wide format. We change the column names of this matrix, and append the lon-lat coordinates to it as follows.

PC

EOF

```
V <- E$v
colnames(E$v) <- paste0("EOF", 1:ncol(SSTdata)) # label columns
EOFs <- cbind(SSTlonlat[-delete_rows, ], E$v)
head(EOFs[, 1:6])

##     lon lat         EOF1         EOF2        EOF3         EOF4
## 16 154 -29 -0.004915064 -0.012129566 -0.02882162 8.540892e-05
## 17 156 -29 -0.001412275 -0.002276177 -0.02552841 6.726077e-03
## 18 158 -29  0.000245909  0.002298082 -0.01933020 8.591251e-03
## 19 160 -29  0.001454972  0.002303585 -0.01905901 1.025538e-02
## 20 162 -29  0.002265778  0.001643138 -0.02251571 1.125295e-02
## 21 164 -29  0.003598762  0.003910823 -0.02311128 1.002285e-02
```

The matrix $\mathbf{U}$ returned from **svd** contains the principal component time series in wide-table format (i.e., each column corresponds to a time series associated with an EOF). Here we use the function **gather** in the package **tidyr** that reverses the operation **spread**. That is, the function takes a spatio-temporal data set in wide-table format and puts it into long-table format. We instruct the function to gather every column except the column denoting time, and we assign the key–value pair EOF-PC:

Now, consider the SVD of the detrended and scaled data matrix,

$$\widetilde{\mathbf{Z}} = \mathbf{U}\mathbf{D}\mathbf{V}', \tag{2.12}$$

where $\mathbf{U}$ is the $T \times T$ matrix of left singular vectors, $\mathbf{D}$ is a $T \times m$ matrix containing singular values on the main diagonal, and $\mathbf{V}$ is an $m \times m$ matrix containing the right singular vectors, where both $\mathbf{U}$ and $\mathbf{V}$ are orthonormal matrices. Upon substituting

(2.12) into (2.11), it is easy to see that the EOFs are given by $\boldsymbol{\Psi} = \mathbf{V}$, and $\boldsymbol{\Lambda} = \mathbf{D}'\mathbf{D}$. In addition, it is straightforward to show that $\mathbf{A} = (\sqrt{T-1})\mathbf{U}\mathbf{D}$ and that the first m columns of $(\sqrt{T-1})\mathbf{U}$ correspond to the normalized PC time series, $\mathbf{A}_{\mathrm{norm}}$. Thus, the advantages of the SVD calculation approach are: (1) we do not need to calculate the empirical spatial covariance matrix; (2) we get the normalized PC time series and EOFs simultaneously; and (3) the procedure still works when $T < m$. The case of $T < m$ can be problematic in the covariance context since then $\mathbf{C}_z^{(0)}$ is not positive-definite, although, as shown in Cressie and Wikle (2011, Section 5.3.4), in this case one can still calculate the EOFs and PC time series.

Fig.2.20(一番上)を描く (p.72-)

```
#作図
g <- ggplot(EOFs, aes(x = lon, y = lat, fill = EOF1))
t <- geom_tile()
s <- fill_scale(name = "degC")
lb <- labs(title = "Empirical Orthogonal Functions", x = "Longitude (deg)", y =
"Latitude (deg)", fill = "degC")
g+t+s+lb+theme_bw()
```

Fig.2.23を描く（p.74-）

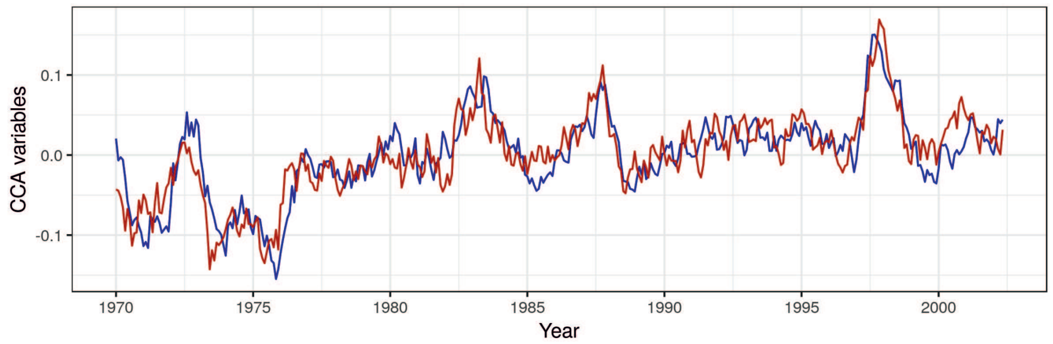


Figure 2.23: Time series of the first canonical variables, $\{a_1, b_1\}$, for $\tau = 7$ month lagged monthly SST anomalies at time $t_j - \tau$ (blue) and those at time t_j (red).

```
nEOF <- 10
#どの式？？
EOFset1 <- E$u[1:(nT-7), 1:nEOF] * sqrt(nT - 1)
EOFset2 <- E$u[8:nT, 1:nEOF] * sqrt(nT - 1)

cc <- cancor(EOFset1, EOFset2)
options(digits = 3)
print(cc$cor[1:5])
print(cc$cor[6:10])

CCA_df <- data.frame(t = 1:(nT - 7),
                     CCAvar1 = (EOFset1 %*% cc$xcoef[,1])[,1],
                     CCAvar2 = (EOFset2 %*% cc$ycoef[,1])[,1])
```

Fig.2.23を描く (p.74-)

```
t_breaks <- seq(1, nT, by = 60)
year_breaks <- seq(1970, 2002, by = 5)

g <- ggplot(CCA_df)
l1 <- geom_line(aes(x = t, y = CCAvar1), colour = "dark blue")
l2 <- geom_line(aes(x = t, y = CCAvar2), colour = "dark red")
s <- scale_x_continuous(breaks = t_breaks, labels = year_breaks)
lb <- labs(title = "Spatio-Temporal CCA", x = "Year", y = "CCA vaiables")
g+l1+l2+s+lb+theme_bw()
```

Fig.2.24を描く (textにコードない; p.75-)

```
EOFs_CCA <- EOFs[,1:4]
EOFs_CCA[,3] <- c(as.matrix(EOFs[,3:12]) %*% cc$xcoef[,1])
EOFs_CCA[,4] <- c(as.matrix(EOFs[,3:12]) %*% cc$ycoef[,1])

f24 <- EOFs_CCA %>%
      gather(key = EOF, value = CCA, EOF1, EOF2) #meltと同じ
g <- ggplot(f24, aes(x = lon, y = lat, fill = CCA))
t <- geom_tile()
s <- fill_scale(name = "degC")
f <- facet_wrap(~ EOF, ncol = 2)
lb <- labs(title = "Spatial weights maps", x = "Longitude (deg)", y = "Latitude (deg)",
fill = "degC")
g+t+s+f+lb+theme_bw()
```

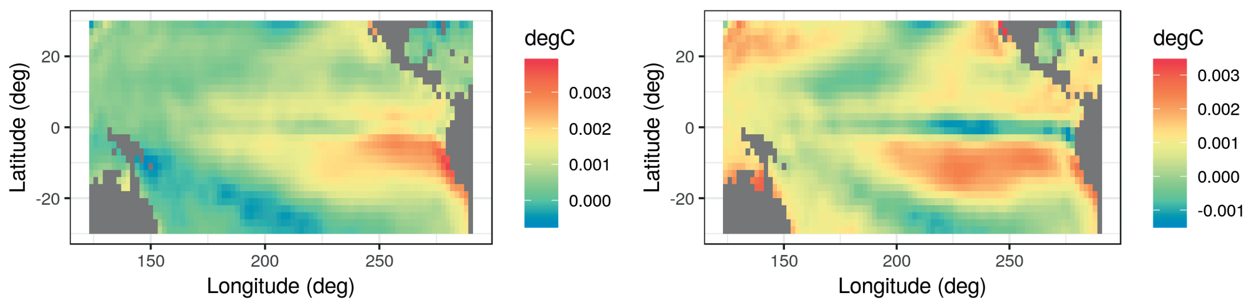


Figure 2.24: Spatial-weights maps corresponding to the linear combination of EOFs used to construct the canonical variables for SST data lagged $\tau = 7$ months (left) and the unlagged SST data (right).