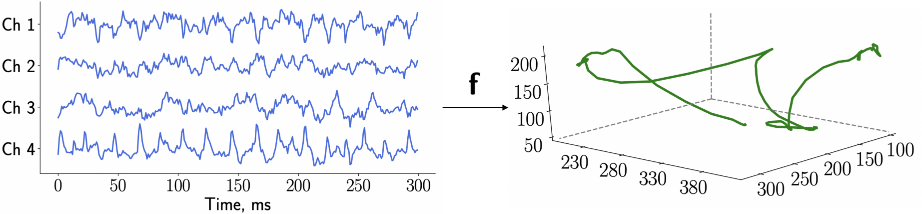
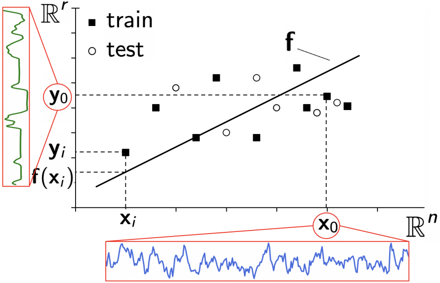
Восстановление зависимости в исходном и целевом пространствах



Прогностическая модель декодирования



Согласование зависимостей в скрытом пространстве

$$\begin{tikzcd} \mathbf{x} \in \mathbb{R}^n \arrow[rr, "\mathbf{f}"] \arrow[dr, bend right, "\mathbf{W}"'] & & \mathbf{y} \in \mathbb{R}^r \arrow[dl, bend left, "\mathbf{C}"] \\ & \mathbf{t}, \mathbf{u} \in \mathbb{R}^\ell \arrow[ul, bend right, "\mathbf{P}"'] \arrow[ur, bend left, "\mathbf{Q}"] & \end{tikzcd}$$

$$\mathbf{x} = \mathbf{P}\mathbf{t} + \mathbf{e}_\mathbf{x}$$

$$\mathbf{y} = \mathbf{Q}\mathbf{u} + \mathbf{e}_\mathbf{y}$$

$$\text{cov}(\mathbf{t}, \mathbf{u}) \to \max_{\mathbf{P}, \mathbf{Q}}$$

Задача декодирования сигналов

$\mathbf{Y} = \mathbf{F}(\mathbf{X}, \boldsymbol{\Theta}) + \mathbf{E}_\mathbf{y} = \mathbf{X} \cdot \boldsymbol{\Theta}^\mathsf{T} + \mathbf{E}_\mathbf{y}$ – модель с параметрами $\boldsymbol{\Theta} \in \mathbb{R}^{r \times n}$.

Функция потерь модели декодирования

$$\mathcal{L}(f, \mathbf{X}, \mathbf{Y}) = \|\mathbf{Y} - \mathbf{F}(\mathbf{X}, \boldsymbol{\Theta})\|_2^2 = \Bigg\| \underset{m \times r}{\mathbf{Y}} - \underset{m \times n}{\mathbf{X}} \cdot \underset{r \times n}{\boldsymbol{\Theta}}^\mathsf{T} \Bigg\|_2^2 \to \min_{\boldsymbol{\Theta}}.$$

Метод проекции в скрытое пространство

$$\mathbf{X} = \mathbf{T}\mathbf{P}^\mathsf{T} + \mathbf{E}_\mathbf{x},$$

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}^\mathsf{T} + \mathbf{E}_\mathbf{y}.$$

$$\begin{tikzcd}
\mathbb{X} \subset \mathbb{R}^n \arrow[rr, "\mathbf{f}"] \arrow[d, shift right, "\mathbf{W}"'] & & \mathbb{Y} \subset \mathbb{R}^r \arrow[d, shift left, "\mathbf{C}"] \\
\mathbb{T} \subset \mathbb{R}^\ell \arrow[rr, "\mathbf{B}"] \arrow[u, shift right, "\mathbf{P}"'] & & \mathbb{U} \subset \mathbb{R}^s \arrow[u, shift left, "\mathbf{Q}"]
\end{tikzcd}$$

Согласование проекций

Для получения согласованной модели в скрытом пространстве находится функция связи

$$\mathbf{U} = \mathbf{T}\mathbf{B}, \quad \mathbf{B} = \operatorname{diag}(\beta_k), \quad \beta_k = \boldsymbol{\nu}_k^\mathsf{T} \boldsymbol{\tau}_k / (\boldsymbol{\tau}_k^\mathsf{T} \boldsymbol{\tau}_k).$$

Финальная модель декодирования имеет вид

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}^\mathsf{T} + \mathbf{E}_\mathbf{y} \approx \mathbf{T}\mathbf{B}\mathbf{Q}^\mathsf{T} + \mathbf{E}_\mathbf{y} = \mathbf{X}\mathbf{W}^*\mathbf{B}\mathbf{Q}^\mathsf{T} + \mathbf{E} = \mathbf{X}\boldsymbol{\Theta}^\mathsf{T} + \mathbf{E}_\mathbf{y},$$

$$\mathbf{T} = \mathbf{X}\mathbf{W}^*, \quad \text{где } \mathbf{W}^* = \mathbf{W}(\mathbf{P}^\mathsf{T}\mathbf{W})^{-1}.$$

Согласование зависимостей в задаче декодирования

Особенностью задачи является избыточность размерности пространств переменных $\mathbf{x}$ и $\mathbf{y}$. Требуется найти многообразия низкой размерности:

$$\begin{tikzcd}
\mathbb{X} \subset \mathbb{R}^n \arrow[r, "\mathbf{f}"] \arrow[d, shift right, "\boldsymbol{\varphi}_\mathbf{x}"'] & \mathbb{Y} \subset \mathbb{R}^r \arrow[d, shift left, "\boldsymbol{\varphi}_\mathbf{y}"] \\
\mathbb{T} \subset \mathbb{R}^\ell \arrow[r, "\mathbf{h}"] \arrow[u, shift right, "\boldsymbol{\psi}_\mathbf{x}"'] & \mathbb{U} \subset \mathbb{R}^s \arrow[u, shift left, "\boldsymbol{\psi}_\mathbf{y}"]
\end{tikzcd}$$

$\mathbb{T} \subset \mathbb{R}^\ell$ и $\mathbb{U} \subset \mathbb{R}^s$ *скрытые пространства* для $\mathbb{X} \in \mathbb{R}^n$ $(\ell \leqslant n)$ и $\mathbb{Y} \in \mathbb{R}^r (s \leqslant r)$, если существуют *функции кодирования* $\boldsymbol{\varphi}_\mathbf{x} : \mathbb{X} \to \mathbb{T}$, $\boldsymbol{\varphi}_\mathbf{y} : \mathbb{Y} \to \mathbb{U}$ и *декодирования* $\boldsymbol{\psi}_\mathbf{x} : \mathbb{T} \to \mathbb{X}$, $\boldsymbol{\psi}_\mathbf{y} : \mathbb{U} \to \mathbb{Y}$:

$$\text{для любого } \mathbf{x} \in \mathbb{X} \quad \text{существует } \mathbf{t} \in \mathbb{T} : \boldsymbol{\psi}_\mathbf{x}\bigl(\boldsymbol{\varphi}_\mathbf{x}(\mathbf{x})\bigr) = \boldsymbol{\psi}_\mathbf{x}(\mathbf{t}) = \mathbf{x};$$

$$\text{для любого } \mathbf{y} \in \mathbb{Y} \quad \text{существует } \mathbf{u} \in \mathbb{U} : \boldsymbol{\psi}_\mathbf{y}\bigl(\boldsymbol{\varphi}_e(\mathbf{y})\bigr) = \boldsymbol{\psi}_\mathbf{y}(\mathbf{u}) = \mathbf{y}.$$

Скрытые пространства $\mathbb{T}$ и $\mathbb{U}$ называются **согласованными**, если существует *функция связи* $\mathbf{h} : \mathbb{T} \to \mathbb{U}$:

$$\mathbf{y} = \mathbf{f}(\mathbf{x}) = \boldsymbol{\psi}_\mathbf{y}\Bigl(\mathbf{h}\bigl(\boldsymbol{\varphi}_\mathbf{x}(\mathbf{x})\bigr)\Bigr).$$

Функция согласования проекций

$$g : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}, \quad g(\boldsymbol{\tau}, \boldsymbol{\nu}) \to \max_{\boldsymbol{\varphi}_x, \boldsymbol{\varphi}_y, \mathbf{h}}$$

Lab Sci III

Granger, causality

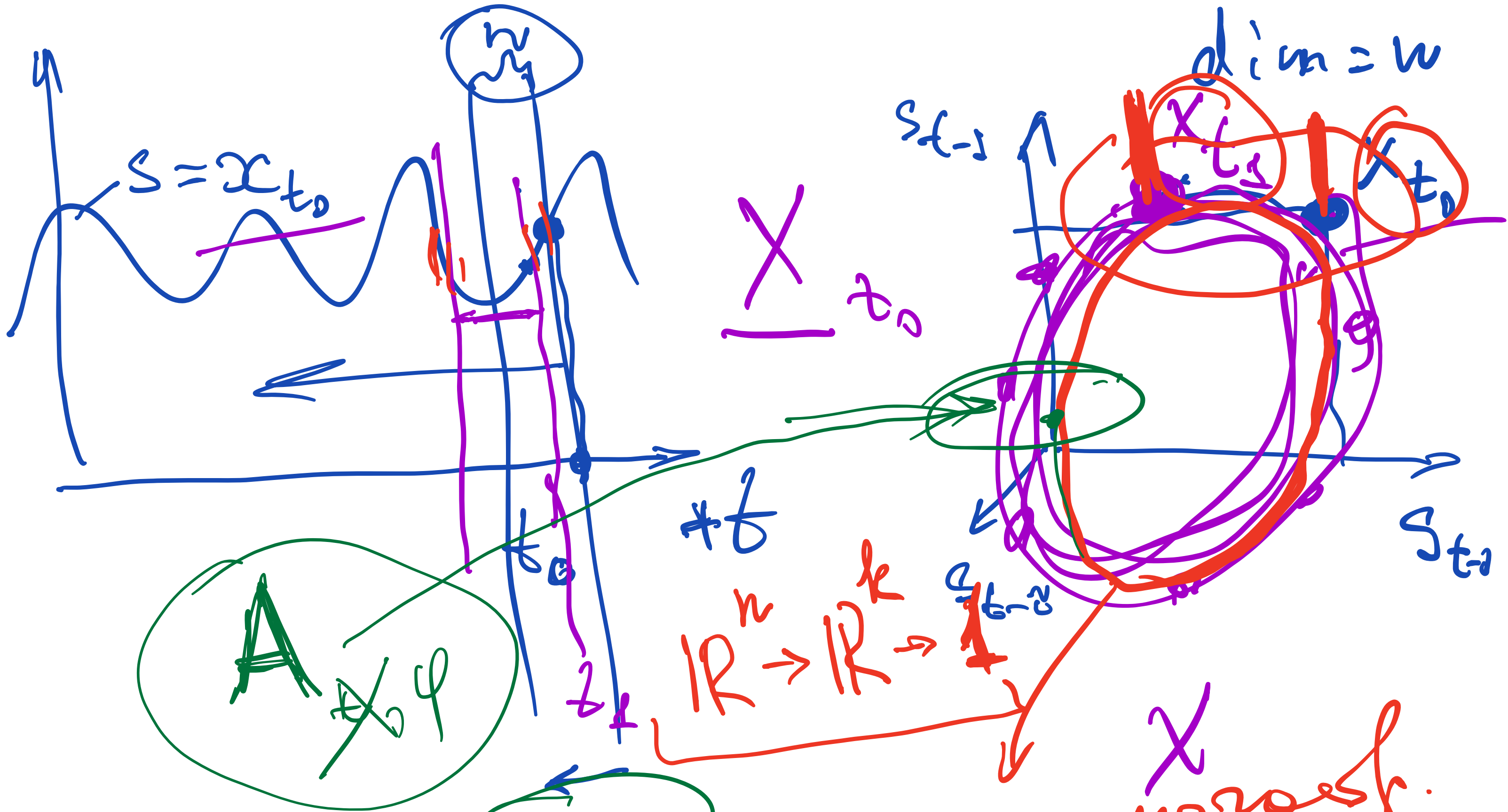
CCM, relation Sugihara

DTW, DBA, time-alignment

SEMOR, scale-alignment

PLS, dimensionality alignment

HOPLS, index alignment



Periodic.

комп. многообр. разм = 1.

Согласованная модель проекции в скрытое пространство

Утверждение (Исаченко, 2017)

Вычисленные вектора $\boldsymbol{\tau}_k$ *и* $\boldsymbol{\nu}_k$ *с помощью итеративной процедуры обновления:*

$$\boldsymbol{\tau}_k := \frac{\mathbf{X}_k \mathbf{w}_k}{\|\mathbf{w}_k\|}, \quad \mathbf{w}_k := \mathbf{X}_k^{\mathsf{T}} \boldsymbol{\nu}_{k-1} / (\boldsymbol{\nu}_{k-1}^{\mathsf{T}} \boldsymbol{\nu}_{k-1});$$

$$\boldsymbol{\nu}_k := \frac{\mathbf{Y}_k \mathbf{c}_k}{\|\mathbf{c}_k\|}, \quad \mathbf{c}_k := \mathbf{Y}_k^{\mathsf{T}} \boldsymbol{\tau}_k / (\boldsymbol{\tau}_k^{\mathsf{T}} \boldsymbol{\tau}_k).$$

обладают максимальной ковариацией $\mathrm{cov}(\boldsymbol{\tau}, \boldsymbol{\nu})$.

Теорема (Исаченко, 2017)

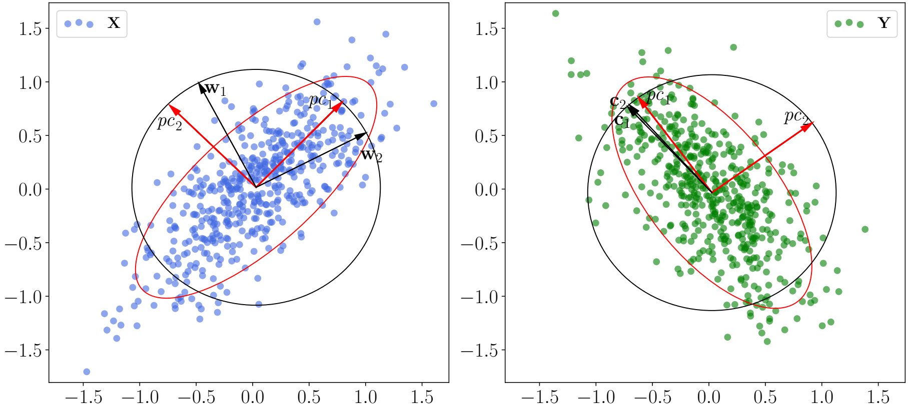
В случае линейных функций декодирования $\varphi_{\mathbf{x}}(\mathbf{T}) = \mathbf{T}\mathbf{P}^{\mathsf{T}}$, $\varphi_{\mathbf{y}}(\mathbf{U}) = \mathbf{U}\mathbf{Q}^{\mathsf{T}}$ *и функции согласования* $g(\boldsymbol{\tau}, \boldsymbol{\nu}) = \mathrm{cov}(\boldsymbol{\tau}, \boldsymbol{\nu})$ *параметры*

$$\boldsymbol{\Theta} = \mathbf{W}(\mathbf{P}^{\mathsf{T}}\mathbf{W})^{-1}\mathbf{B}\mathbf{Q}^{\mathsf{T}}$$

являются оптимальными для модели $\mathbf{F}(\mathbf{X}, \boldsymbol{\Theta})$.

Пример согласованной проекции в скрытое пространство

Исходные переменные $\mathbf{x}_i \sim \mathcal{N}(0, \mathbf{\Sigma})$.
Целевые переменные $\mathbf{y}_i$ линейно зависят от pc_2 и не зависят от pc_1.



Согласование проекций матриц $\mathbf{X}$ и $\mathbf{Y}$ находит оптимальное скрытое представление, отклоняя вектора $\mathbf{w}_k$ и $\mathbf{c}_k$ от направления главных компонент.

Композиция моделей декодирования сигналов

Пусть $\mathbf{f}_1(\mathbf{x}_1, \boldsymbol{\Theta}_1)$, $\mathbf{f}_2(\mathbf{x}_2, \boldsymbol{\Theta}_2)$ – линейные модели декодирования сигналов.

Утверждение (Исаченко, 2021)

Пусть модель декодирования является аддитивной композицией линейных моделей:

$$\mathbf{y} = \mathbf{f}_1(\mathbf{x}_1, \boldsymbol{\Theta}_1) + \mathbf{f}_2(\mathbf{x}_2, \boldsymbol{\Theta}_2) + \varepsilon_{\mathbf{y}} = \boldsymbol{\Theta}_1\mathbf{x}_1 + \boldsymbol{\Theta}_2\mathbf{x}_2 + \varepsilon_{\mathbf{y}}.$$

Тогда оптимальные параметры

$$\boldsymbol{\Theta}_1 = (\mathbf{X}_1^{\mathsf{T}}\mathbf{M}_{\mathbf{X}_2}\mathbf{X}_1)^{-1}\mathbf{X}_1^{\mathsf{T}}\mathbf{M}_{\mathbf{X}_2}\mathbf{Y}, \quad \boldsymbol{\Theta}_2 = (\mathbf{X}_2^{\mathsf{T}}\mathbf{M}_{\mathbf{X}_1}\mathbf{X}_2)^{-1}\mathbf{X}_2^{\mathsf{T}}\mathbf{M}_{\mathbf{X}_1}\mathbf{Y},$$

где

$$\mathbf{M}_{\mathbf{X}_1} = \mathbf{I} - \mathbf{P}_{\mathbf{X}_1},\ \mathbf{P}_{\mathbf{X}_1} = \mathbf{X}_1(\mathbf{X}_1^{\mathsf{T}}\mathbf{X}_1)^{-1}\mathbf{X}_1^{\mathsf{T}},$$
$$\mathbf{M}_{\mathbf{X}_2} = \mathbf{I} - \mathbf{P}_{\mathbf{X}_2},\ \mathbf{P}_{\mathbf{X}_2} = \mathbf{X}_2(\mathbf{X}_2^{\mathsf{T}}\mathbf{X}_2)^{-1}\mathbf{X}_2^{\mathsf{T}}.$$

Теорема (Исаченко, 2021)

Пусть выполнены следующие условия

$$\mathbf{Y} \neq \mathbf{P}_{\mathbf{X}_2}\mathbf{Y}, \quad \mathbf{X}_1 \neq \mathbf{P}_{\mathbf{X}_2}\mathbf{X}_1, \quad \mathbf{Y}^{\mathsf{T}}\mathbf{M}_{\mathbf{X}_2}\mathbf{X}_1 \neq \mathbf{0}.$$

Тогда выполнено строгое неравенство

$$\mathcal{L}_{\mathsf{sup}}(\boldsymbol{\Theta}_1^*, \boldsymbol{\Theta}_2^*, \mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}) < \mathcal{L}(\boldsymbol{\Theta}_2, \mathbf{X}_2, \mathbf{Y}).$$

Нелинейные методы согласования скрытого пространства

Функции кодирования и декодирования являются глубокими нейросетями:

$$\mathbf{T} = \varphi_{\mathbf{x}}(\mathbf{X}) = \mathbf{W}_x^L \sigma(\dots \sigma(\mathbf{W}_x^2 \sigma(\mathbf{X}\mathbf{W}_x^1)) \dots)$$
$$\mathbf{U} = \varphi_{\mathbf{y}}(\mathbf{Y}) = \mathbf{W}_y^L \sigma(\dots \sigma(\mathbf{W}_y^2 \sigma(\mathbf{Y}\mathbf{W}_y^1)) \dots)$$
$$\mathbf{X} = \psi_{\mathbf{x}}(\mathbf{T}) = \mathbf{W}_t^L \sigma(\dots \sigma(\mathbf{W}_x^2 \sigma(\mathbf{T}\mathbf{W}_t^1)) \dots)$$
$$\mathbf{Y} = \psi_{\mathbf{y}}(\mathbf{U}) = \mathbf{W}_u^L \sigma(\dots \sigma(\mathbf{W}_y^2 \sigma(\mathbf{U}\mathbf{W}_u^1)) \dots)$$

Согласование проекций

$$g(\mathbf{T}, \mathbf{U}) \to \max_{\mathbf{W}}, \quad \mathbf{W} = \{\mathbf{W}_x^i, \mathbf{W}_y^i, \mathbf{W}_t^i, \mathbf{W}_u^i\}_{i=1}^L.$$

Градиент функции согласования $g(\boldsymbol{\tau}, \boldsymbol{\nu}) = \text{corr}(\boldsymbol{\tau}, \boldsymbol{\nu})$ имеет вид

$$\frac{\partial g(\boldsymbol{\tau}, \boldsymbol{\nu})}{\partial \boldsymbol{\tau}} = \frac{1}{\ell - 1} \left(\boldsymbol{\Sigma}_1^{-1/2} \mathbf{U}\mathbf{V}^{\mathsf{T}} \boldsymbol{\Sigma}_2^{-1/2} \mathbf{U} - \boldsymbol{\Sigma}_1^{-1/2} \mathbf{U}\mathbf{D}\mathbf{V}^{\mathsf{T}} \boldsymbol{\Sigma}_1^{-1/2} \right),$$

где $\mathbf{U}, \mathbf{D}, \mathbf{V} = \text{SVD}(\boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_1^{-1/2} \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_2^{-1/2}$, $\boldsymbol{\Sigma}_1 = \frac{1}{\ell-1}\mathbf{T}\mathbf{T}^{\mathsf{T}}$, $\boldsymbol{\Sigma}_2 = \frac{1}{\ell-1}\mathbf{U}\mathbf{U}^{\mathsf{T}}$, $\boldsymbol{\Sigma}_{12} = \frac{1}{\ell-1}\mathbf{T}\mathbf{U}^{\mathsf{T}}$.