حلّ مسألهی کاهش بُعد غیرخطّی تفسیرپذیر

ارائه دهنده:

محمّدرضا رحمانی بهراد منیری

استاد:

جناب آقای دکتر مدّاح‌علی

دانشگاه صنعتی شریف

بهمن ۱۳۹۸

چشم‌انداز

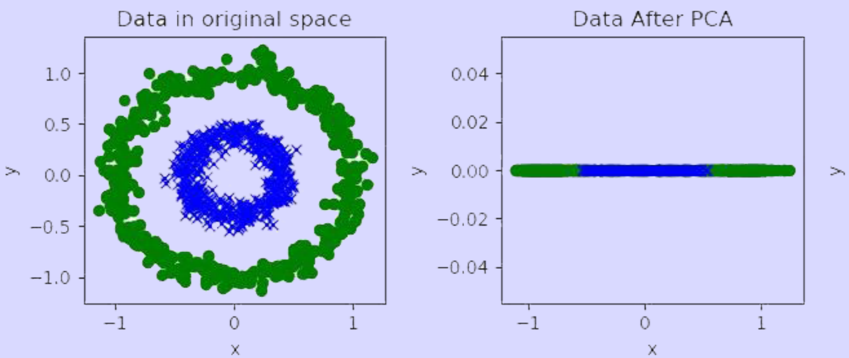
۱ انگیزه

۱.۱ کاهش بعد تفسیرپذیر

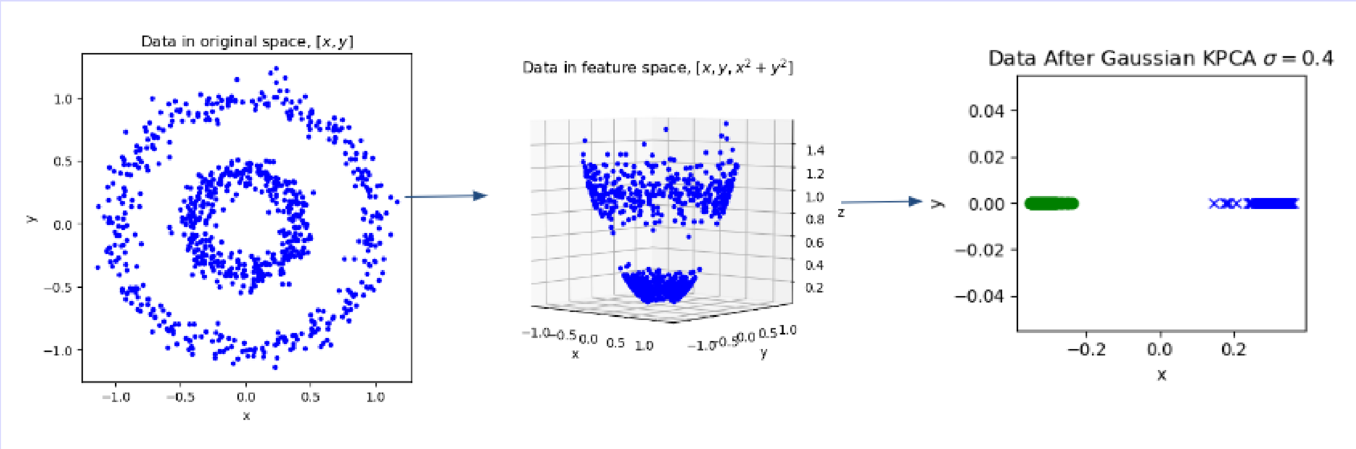
- روش PCA
  - مزیّت: تفسیرپذیری

$$\begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} w_{11}x1 + w_{12}x_2 + w_{13}x_3 \\ w_{21}x_1 + w_{22}x_2 + w_{23}x_3 \end{bmatrix}$$

  - مشکل: عدم توانایی در شناسایی روابط غیرخطّی میان ویژگی‌ها



- روش KPCA
  - مزیّت: شناسایی روابط غیرخطّی میان ویژگی‌ها



  - مشکل ۱: تفسیرناپذیری

$$\Phi(X)W \qquad\qquad \Phi(XW)$$

  - مشکل ۲: عدم استفاده از برچسب داده‌ها در کاهش بُعد

انگیزه
ISM
کاربردها
شبیه‌سازی
مراجع

- IKDR
  - لزوم وجود یک معیار وابستگی

$$\max_{W} \ \mathrm{DEP}(XW, Y) \qquad \text{s.t. } W^\top W = I$$

  - معیار وابستگی HSIC

$$\mathrm{MMD}(P, Q) = \left\| \mathbb{E}_P \left[k(., X) \right] - \mathbb{E}_Q \left[k(., X) \right] \right\|_{\mathcal{H}}$$

$$\mathrm{HSIC}_k(X, Y) \triangleq \mathrm{MMD}_k(P_X P_Y, P_{XY})$$

انگیزه
ISM
کاربردها
شبیه‌سازی
مراجع

- IKDR
  - لزوم وجود یک معیار وابستگی

$$\max_{W} \ \mathrm{DEP}(XW, Y) \qquad \text{s.t. } W^\top W = I$$

  - معیار وابستگی HSIC

$$\mathrm{MMD}(P, Q) = \left\| \mathbb{E}_P\left[k(.,X)\right] - \mathbb{E}_Q\left[k(.,X)\right] \right\|_{\mathcal{H}}$$

$$\mathrm{HSIC}_k(X, Y) \triangleq \mathrm{MMD}_k(P_X P_Y, P_{XY})$$

  - تخمین HSIC از روی داده‌ها

$$\begin{aligned} \mathrm{HSIC}(XW, Y) &= \frac{1}{(n-1)^2}\mathrm{tr}(K_{XW} H K_Y H) \\ &= \frac{1}{(n-1)^2}\mathrm{tr}(H K_Y H K_{XW}) \\ &= \frac{1}{(n-1)^2}\mathrm{tr}(\Gamma K_{XW}) \end{aligned}$$

- حلّ مسأله‌ی IKDR
  - یک مسأله‌ی بهینه‌سازی غیرمحدّب و به شدّت غیرخطّی
  - راه‌حل‌های قبلی موجود برای حلّ این مسأله
    * Stiefel Manifold
    * Grassman Manifold
    * Dimension Growth
  - بهترین راه‌حل: الگوریتم Iterative Spectral Method (ISM)
    * پیاده‌سازی ساده‌تر
    * بسیار سریع‌تر از روش‌های قبلی
    * دچار مشکل نشدن در نقاط زینی
    * تضمین تئوری فقط برای کرنل گاوسی
  - هدف مقاله: تعمیم الگوریتم ISM به خانواده‌ی وسیع‌تری از کرنل‌ها

۲ Iterative Spectral Method

۱.۲ ایده‌ی الگوریتم ISM

تعریف کرنل ISM

کرنل متقارن و مثبت معین $k(.,.)$، یک کرنل ISM است، اگر دو بار مشتق‌پذیر باشد و برای آن داشته باشیم:

$$\forall \mathbf{x}_i,\ \mathbf{x}_j \in \mathbb{R}^d \quad k(\mathbf{x}_i W, \mathbf{x}_j W) = f(\beta_{ij}) \tag{۱}$$

که در آن

$$\beta_{ij} = a(\mathbf{x}_i, \mathbf{x}_j)^\top W W^\top b(\mathbf{x}_i, \mathbf{x}_j). \tag{۲}$$

- مسأله‌ی بهینه‌سازی الگوریتم IKDR :

$$\max_W \ \ \mathrm{tr}(\Gamma K_{XW}) \quad \text{s.t.} \quad W^\top W = I$$

مسأله‌ی بهینه‌سازی الگوریتم IKDR: $\max_W \ \mathrm{tr}(\Gamma K_{XW}) \ \text{ s.t. } \ W^\top W = I$

قضیه

اگر

$$-\frac{1}{2}\sum_{ij}\Gamma_{ij}f'(\mathbf{a}^T W W^T \mathbf{b})(\mathbf{b}\mathbf{a}^\top + \mathbf{a}\mathbf{b}^\top)W - W\Lambda = 0.$$

باشد، شرایط مرتبه‌ی اوّل مسأله‌ی بهینه‌سازی IKDR برقرار می‌شود. این به آن معناست که W باید شامل بردارهای ویژه‌ی ماتریس

$$\Phi = -\frac{1}{2}\sum_{ij}\Gamma_{ij}f'(\mathbf{a}^T W W^T \mathbf{b})(\mathbf{b}\mathbf{a}^\top + \mathbf{a}\mathbf{b}^\top)$$

باشد. هم‌چنین اگر W شامل بردارویژه‌های متناظر با کوچک‌ترین مقادیر ویژه‌ی Φ باشد، در شرایط مرتبه‌ی دوم مسأله‌ی بهینه‌سازی IKDR هم صدق می‌کند.

- الگوریتم ISM:

Algorithm 1: ISM Algorithm

Input: Data X, kernel, Subspace Dimension q

Output: Projected subspace W

Initialization: Initialize Φ_0.

Set W_0 to Dominant Eigenvectors of of Φ_0

while $||\Lambda_i - \Lambda_{i-1}||_2/||\Lambda_i||_2 < \delta$ **do**

  Compute Φ with W_{k-1}

  Set W_k to Dominant Eigenvectors of Φ

end

- انتخاب یک شرط اولّیه‌ی مناسب: Φ_0

- الگوریتم ISM:

Algorithm 1: ISM Algorithm

Input: Data X, kernel, Subspace Dimension q

Output: Projected subspace W

Initialization: Initialize Φ_0.

Set W_0 to Dominant Eigenvectors of of Φ_0

while $||\Lambda_i - \Lambda_{i-1}||_2/||\Lambda_i||_2 < \delta$ **do**

  Compute Φ with W_{k-1}

  Set W_k to Dominant Eigenvectors of Φ

end

- انتخاب یک شرط اولّیه‌ی مناسب: Φ_0

قضیه

هر ترکیب خطّی با ضرایب مثبت ازکرنل‌های ISM، خود یک کرنل ISM است و ماتریس Φ آن، از ترکیب خطّی ماتریس‌های Φ کرنل‌های اولیه با همان ضرایب به دست می آید.

۳ کاربردها

۱.۳ کاربردهای الگوریتم ISM

- مسأله‌ی کاهش بُعد با نظارت (Supervised Dimension Reduction)
- مسأله‌ی کاهش بُعد بدون نظارت (Unsupervised Dimension Reduction)
- مسأله‌ی کاهش بُعد با نظارت ناقص (Semi-Supervised Dimension Reduction)
- مسأله‌ی کاهش بعد جایگزین (Alternative Clustering)

۴ شبیه‌سازی

۵ مراجع

مراجع

[1] Navabi, Zainalabedin. *Digital System Test and Testable Design*. Springer, 2011.

[2] Marconi Communications GmbH. *Introduction to the Synchronous Digital Hierarchy*. Marconi Communications GmbH, 3.2 ed. , 2000.

[۳] پرمان. مستند طرّاحی سیستم *PTS6010* . پرمان، ویرایش ۰.۲، ۱۳۹۴.

با تشکّر