

## بررسی الگوهای همبستگی دما، رطوبت و بارندگی در ایران با استفاده از تحلیل همبستگی کانونی

آناهیتا متقی گلشن، سید محمدابراهیم حسینی‌نسب\* و محمدرضا فرید روحانی

دانشگاه شهید بهشتی

چکیده. یکی از مسایل مورد توجه در تحلیل متغیرهای بردار مقدار، بررسی ساختار همبستگی و ارتباط میان دو مجموعه از چنین متغیرهایی است. هنگامی که دو مجموعه داده، هرکدام شامل تعداد زیادی متغیر باشند، محقق می‌تواند ارتباط میان دو مجموعه را با یافتن تعداد محدودی ترکیب خطی از متغیرهای مجموعه‌ی اول که دارای بیش‌ترین همبستگی با ترکیب خطی از متغیرهای مجموعه‌ی دوم است بیان کند. این روش به تحلیل همبستگی کانونی موسوم است.

اگر مشاهدات، توابع پیوسته‌ای از یک متغیر (معمولاً زمان) باشند، نمی‌توان روش‌های موجود برای تحلیل داده‌های چندمتغیره را برای تحلیل آن‌ها به‌کار گرفت. بنا بر این برخی انطباق‌ها در تعاریف و مفاهیم با در نظر گرفتن ماهیت داده‌ها لازم به نظر می‌رسد. این امر به تعریف همبستگی کانونی برای جفت توابع تصادفی منجر می‌شود که به آن همبستگی کانونی تابعی گویند. (FCCA) در این مقاله، داده‌های آب و هوای ایران در سال ۲۰۰۶ با استفاده از FCCA مورد تحلیل قرار گرفته است. مجموعه داده‌های مورد بررسی شامل مقادیر گسسته سه پدیده‌ی دما، رطوبت و بارندگی است که از ۱۰۲ ایستگاه هواشناسی جمع‌آوری شده‌اند. در تحلیل این داده‌ها، ابتدا منحنی‌های مناسب به داده‌های اولیه برازش و سپس الگوهای همبستگی میان هر جفت از سه پدیده شناسایی و استخراج شده است.

واژگان کلیدی. تحلیل داده‌های تابعی؛ همبستگی کانونی تابعی؛ عملگر کوواریانس؛ عملگر هیلبرت اشمیت؛ هموارسازی.

\* نویسنده‌ی عهده‌دار مکاتبات

## ۱ مقدمه

یکی از مسایل مورد توجه در آمار، بررسی ساختار همبستگی و ارتباط میان متغیرهاست. روش‌های مختلفی برای بررسی این ارتباط وجود دارد که از هر کدام بسته به ویژگی‌های داده‌های مورد بررسی و هدف تحقیق استفاده می‌شود. چنانچه دو مجموعه از متغیرها شامل تعداد زیادی متغیر باشند و خلاصه کردن ارتباط میان دو مجموعه در تعداد محدودی ترکیب خطی که دارای برخی خواص ماکسیمال مورد نظر هستند هدف مطالعه باشد از تحلیل همبستگی کانونی استفاده می‌شود. تحلیل همبستگی کانونی متضمن افزاز کردن گردایه‌ای از متغیرها به دو مجموعه‌ی  $x$  (شامل  $p$  متغیر) و  $y$  (شامل  $q$  متغیر) می‌باشد که هدف آن بدست آوردن ترکیب‌های خطی  $u = a^T x$  و  $v = b^T y$  به قسمی است که  $u$  و  $v$  دارای بیش‌ترین همبستگی ممکن باشد.

اگر مشاهدات، توابع پیوسته‌ای از یک متغیر (معمولاً زمان) باشند، نمی‌توان روش‌های موجود برای تحلیل داده‌های چندمتغیره را برای تحلیل آن‌ها بکار گرفت. بنا بر این تحلیل آن‌ها مستلزم ارایه‌ی نظریه‌های لازم با در نظر گرفتن ماهیت داده‌ها بوده و انطباق‌هایی در تئوری‌های موجود لازم به نظر می‌رسد. مطالعه‌ی همبستگی کانونی برای چنین داده‌هایی به یک گرایش جدید در آمار منتهی می‌شود که به آن تحلیل همبستگی کانونی تابعی (FCCA) گویند. در صورتی که مشاهدات حاصل از پدیده‌های تابعی به صورت گسسته جمع‌آوری شده باشند اولین گام در تحلیل آن‌ها، برازش منحنی‌های مناسب به این داده‌های گسسته‌ی اولیه می‌باشد که به آن ثبت‌نام داده‌ها گویند. این اقدام به دلیل تابعی بودن طبیعت پدیده‌های مورد مطالعه است. از طرف دیگر، مشاهدات اولیه ممکن است با خطا اندازه‌گیری شوند، از این‌رو لازم است تا ابتدا با استفاده از یکی از روش‌های هموارسازی، این مشاهدات به توابع پیوسته تبدیل شوند.

تحلیل همبستگی کانونی دارای برخی خاصیت‌های ماکسیمال مشابه خاصیت‌های تحلیل مؤلفه‌ی اصلی است، با این تفاوت که تحلیل مؤلفه‌ی اصلی، همبستگی‌های درونی بین مجموعه‌ای از متغیرها را در نظر می‌گیرد، در حالی‌که، تمرکز همبستگی کانونی بر همبستگی بین دو گروه از متغیرهاست. تحلیل همبستگی کانونی را می‌توان از جنبه‌ی دیگری نیز مورد توجه قرار داد، بدین صورت که این تحلیل را تعمیمی از تحلیل رگرسیونی در نظر گرفت. همان‌طور که می‌دانیم در رگرسیون در پی یافتن ترکیب خطی  $a^T x$  از متغیرهای مستقل  $x$  می‌باشیم که قویاً با  $y$  همبسته باشد، حال آن‌که تحلیل همبستگی کانونی بر یافتن ترکیب‌های خطی  $a^T x$  و  $b^T y$  که قویاً با هم همبسته باشند متمرکز است. بنا بر این دو تحلیل تنها در تعداد متغیرهای مجموعه  $y$  تفاوت دارند؛ به این معنی که در تحلیل همبستگی کانونی تعداد متغیرهای مجموعه‌ی  $y$  بیش از ۱ است.

تحلیل همبستگی کانونی چندمتغیره توسط هتلینگ (۱۹۳۶) به حالت تابعی بسط داده شده است. هانا (۱۹۶۱) و بریلینگر (۱۹۷۵) تحلیل همبستگی کانونی را برای سری‌های زمانی مانای چندمتغیره به کار بردند و با بکارگیری مانایی، مسئله را به تحلیل کانونی چندمتغیره‌ی کلاسیک تبدیل کردند. یک روش نظری که روی زوایای میان زیرفضاها پایه‌ریزی شده توسط داکسیون و اینکیت (۱۹۹۷) بیان شده است. نمونه‌ای از همبستگی کانونی تابعی هموار شده توسط لرگانس و همکاران (۱۹۹۳) تعریف شده است، که در آن به لزوم استفاده از روش‌های هموارساز در تحلیل همبستگی کانونی تابعی اشاره شده است. آن‌ها برای هموارسازی از اسپالین‌ها استفاده نمودند و به‌وسیله‌ی آن مجموعه داده‌های واقعی «قدم زدن انسان» را تحلیل کردند. برای جزئیات بیشتر در مورد مسایل مربوط به بسط مؤلفه‌ی اصلی از حالت چندمتغیره به حالت تابعی می‌توان به رایس و سیلورمن (۱۹۹۱) مراجعه نمود. سرویس و همکاران (۱۹۹۸) نیز برای تحلیل زوج توابع تصادفی با استفاده از توابع کوواریانس متقاطع اقدام نمودند. استفاده از هموارسازی در تحلیل همبستگی کانونی موجب کاهش بُعد می‌شود که این می‌تواند با انتخاب مناسب تاوان برای اسپالین‌های هموارساز بدست آید (لرگانس و همکاران، ۱۹۹۳). همچنین این اقدام به‌وسیله‌ی روش‌های دیگری که از مشکل وارون جلوگیری می‌کند امکان‌پذیر است (سرویس و همکاران، ۱۹۹۸). روشی که در این مقاله به آن پرداخته شده است به برآورد توابع مشاهده‌شده به‌وسیله‌ی بسط‌های متناهی برحسب توابع ویژه می‌پردازد و سپس تحلیل همبستگی کانونی را در یک زیر فضای متناهی انجام می‌دهد (رمسی و سیلورمن، ۲۰۰۵). این مقاله به تحلیل همبستگی کانونی در داده‌های تابعی همراه با تحلیل داده‌های مربوط به آب و هوای ایران در سال ۲۰۰۶ پرداخته است.

## ۲ تحلیل همبستگی کانونی تابعی

داده‌های چندمتغیره بر اساس ماتریس مشاهدات تحلیل می‌شوند که در آن مؤلفه‌ی  $x_{ij}$  نشان‌دهنده‌ی مقادیر متغیر  $z$ ام برای آزمودنی  $i$ ام می‌باشد. اما در تحلیل داده‌های تابعی، مقدار متغیر برای آزمودنی  $i$ ام به صورت  $x_i(t)$  است که در آن  $t \in I = [0, T]$ . با وجود این تفاوت، با اقتباس از نگاشت‌های ماتریسی در حالت چندمتغیره، نگاشت‌های تابعی برای مشاهداتی با بُعد بی‌نهایت نیز قابل تعریف هستند (رمسی، ۱۹۸۲).  
تعریف. فرایند دومتغیره‌ی  $(X, Y)$  که در آن  $X(\cdot) \in L_2(T_1)$  و  $Y(\cdot) \in L_2(T_2)$  و دارای یک توزیع توأم می‌باشند را در نظر بگیرید.

$$\int E \|Z\|^2 = E[\langle ZZ \rangle] = E \int_T (Z(s))^2 ds < \infty \quad Z = X \text{ یا } y$$

که در آن  $T_1$  و  $T_2$  دو مجموعه‌ی اندیس‌گذار و  $L_2(T_1)$  و  $L_2(T_2)$  دو فضای هیلبرت هستند که توابع موجود در آن‌ها به‌ترتیب از  $T_1$  و  $T_2$  به  $R$  تعریف شده‌اند و انتگرال توان دوم آن‌ها متناهی است. اندازه‌های متناظر با فضاها  $L_2(T_1)$  و  $L_2(T_2)$  را به‌ترتیب با  $\mu_1$  و  $\mu_2$  نشان داده و حاصل ضرب داخلی مرتبط با هر فضا به‌صورت زیر تعریف می‌شود:

$$\langle u, v \rangle = \int u(s)v(s)d\mu_i(s) \quad i = 1, 2.$$

اندازه‌های  $\mu_1$  و  $\mu_2$  معمولاً اندازه‌ی لیگ اختیار می‌شوند.

همبستگی کانونی در حالت چندمتغیره برای بردارهای بُعد متناهی  $X \in R^{k_1}$  و  $Y \in R^{k_2}$  و در حالت تابعی برای فرایندهای تصادفی  $X(\cdot) \in L_2(T_1)$  و  $Y(\cdot) \in L_2(T_2)$  تعریف می‌شود. در این صورت اولین همبستگی کانونی  $\rho_1$  و توابع یا بردارهای وزنی مرتبط با آن،  $u_1(\cdot)$  و  $v_1(\cdot)$  به‌صورت زیر بدست می‌آیند:

$$\begin{aligned} \rho_1 &= \sup_{\substack{u \in H_1, \text{var}(\langle u, X \rangle) = 1, \\ v \in H_2, \text{var}(\langle v, Y \rangle) = 1}} \text{cov}(\langle u, X \rangle, \langle v, Y \rangle) \\ &= \text{cov}(\langle u_1, X \rangle, \langle v_1, Y \rangle), \end{aligned}$$

که در آن  $H_i$  در حالت چندمتغیره به‌صورت  $H_i = R^{k_i}$  و در حالت تابعی به‌صورت  $H_i = L_2(T_i)$  تعریف می‌شود.

$k$ امین ( $k > 1$ ) همبستگی کانونی و توابع وزن مربوط به آن که به ترتیب آن‌ها را با  $\rho_k$ ،  $u_k(\cdot)$  و  $v_k(\cdot)$  نشان می‌دهیم به‌صورت زیر تعریف می‌شود:

$$\begin{aligned} \rho_k &= \sup_{\substack{u \in H_1, \text{var}(\langle u, X \rangle) = 1, \\ v \in H_2, \text{var}(\langle v, Y \rangle) = 1}} \text{cov}(\langle u, X \rangle, \langle v, Y \rangle) \\ &= \text{cov}(\langle u_k, X \rangle, \langle v_k, Y \rangle), \end{aligned}$$

که در آن  $k$ امین جفت از متغیرهای کانونی  $(U_k, V_k)$  با  $k-1$  جفت  $\{(U_i, V_i), i = 1, 2, \dots, k-1\}$  نا همبسته است و  $U_k = \langle u_k, X \rangle$  و  $V_k = \langle v_k, Y \rangle$ .

لازم به ذکر است که با وجود تفاوت‌های اساسی در مفهوم و بحث وجودی آن، تعریف همبستگی کانونی در حالت تابعی مشابه با حالت چندمتغیره است. این تشابه با تطابق حاصل ضرب داخلی در حالت

چندمتغیره که برای دو بردار تصادفی  $x$  و  $y$  به صورت  $\langle x, y \rangle = x^T y$  است به حاصل ضرب داخلی در فضاهای هیلبرت مرتبط با دو فرایند تصادفی  $X(\cdot)$  و  $Y(\cdot)$  یعنی  $\langle X, Y \rangle = \int X(t)Y(t)dt$  نمود پیدا می‌کند.

**تعریف.** برای فرایندهای تصادفی  $X \in L_2(T_1)$  و  $Y \in L_2(T_2)$ ، توابع کوواریانس به صورت زیر تعریف می‌شود:

$$r_{XX}(s, t) = \text{cov}\{X(s), X(t)\}, \quad s, t \in T_1,$$

$$r_{YY}(s, t) = \text{cov}\{Y(s), Y(t)\}, \quad s, t \in T_2,$$

$$r_{XY}(s, t) = \text{cov}\{X(s), Y(t)\}, \quad s \in T_1, t \in T_2.$$

همچنین عملگرهای کوواریانس  $R_{XX} : L_2(T_1) \rightarrow L_2(T_1)$ ،  $R_{YY} : L_2(T_2) \rightarrow L_2(T_2)$  و  $R_{XY} : L_2(T_2) \rightarrow L_2(T_1)$  به صورت زیر تعریف می‌شوند:

$$R_{XX}u(s) = \int_{T_1} r_{XX}(s, t)u(t)dt, \quad u \in L_2(T_1),$$

$$R_{XY}u(s) = \int_{T_2} r_{XY}(s, t)u(t)dt, \quad u \in L_2(T_2), s \in T_1,$$

$$R_{YY}u(s) = \int_{T_2} r_{YY}(s, t)u(t)dt, \quad u \in L_2(T_2).$$

لازم به ذکر است که عملگرهای  $R_{XX}$  و  $R_{YY}$  فشرده، هیلبرت اشمیت، متقارن و معین نامنفی و عملگر  $R_{XY}$  فشرده و هیلبرت اشمیت است.

به سادگی می‌توان نشان داد

$$\begin{aligned} \text{cov}(\langle u, X \rangle, \langle v, Y \rangle) &= E\{[\langle u, X \rangle - E(\langle u, X \rangle)] \\ &\quad \times [\langle v, Y \rangle - E(\langle v, Y \rangle)]\} \\ &= E\{[\langle u, X - EX \rangle][\langle v, Y - EY \rangle]\} \\ &= \langle u, R_{XY}v \rangle. \end{aligned}$$

همچنین به طور مشابه داریم:

$$\text{var}(\langle u, X \rangle) = \langle u, R_{XX}u \rangle, \quad \text{var}(\langle v, Y \rangle) = \langle v, R_{YY}v \rangle.$$

با توجه به ویژگی‌ها و خواص عملگرهای کوواریانس همچنین رابطه‌های قبل می‌توان  $k$  امین همبستگی کانونی را به صورت زیر محاسبه کرد:

$$\begin{aligned}\rho_k &= \sup_{\substack{u \in L_2(T_1), \langle u, R_{XX}u \rangle = 1, \\ v \in L_2(T_2), \langle v, R_{YY}v \rangle = 1}} \langle u, R_{YX}v \rangle \\ &= \langle u_k, R_{YX}v_k \rangle,\end{aligned}$$

که در آن برای  $k > 1$ ،  $(U_k, V_k)$  با  $(U_i, V_i)$  برای  $i = 1, 2, \dots, k-1$  ناهمبسته است. برای جزییات بیش‌تر به هی و همکاران (۲۰۰۳) مراجعه نمایید.

## ۲/۱ برآوردهای تجربی

اگر مشاهدات به صورت زوج‌های  $(X_1(t), Y_1(t)), \dots, (X_n(t), Y_n(t))$  باشند؛ آنگاه توابع کوواریانس نمونه‌ای به صورت زیر تعریف می‌شوند:

$$\widehat{\text{cov}}\{X(t_1), X(t_2)\} = \frac{1}{n-1} \sum_{i=1}^n \{X_i(t_1) - \bar{X}(t_1)\} \{X_i(t_2) - \bar{X}(t_2)\},$$

$$\widehat{\text{cov}}\{Y(t_1), Y(t_2)\} = \frac{1}{n-1} \sum_{i=1}^n \{Y_i(t_1) - \bar{Y}(t_1)\} \{Y_i(t_2) - \bar{Y}(t_2)\}.$$

همچنین تابع کوواریانس متقاطع<sup>۱</sup> این دو متغیر به صورت

$$\widehat{\text{cov}}\{X(t_1), Y(t_2)\} = \frac{1}{n-1} \sum_{i=1}^n \{X_i(t_1) - \bar{X}_i(t_1)\} \{Y_i(t_2) - \bar{Y}(t_2)\},$$

و تابع همبستگی متقاطع<sup>۲</sup> آن‌ها به شکل

$$\hat{r}_{XY}(t_1, t_2) = \widehat{\text{corr}}\{X(t_1), Y(t_2)\} = \frac{\widehat{\text{cov}}\{X(t_1), Y(t_2)\}}{\sqrt{\widehat{\text{var}}\{X(t_1)\} \widehat{\text{var}}\{Y(t_2)\}}},$$

تعریف می‌شود.

متغیرهای کانونی نمونه‌ای به وسیله‌ی جایگذاری عملگرهای کوواریانس با نسخه‌های نمونه‌ای آن و

برآورد تجربی عملگرهای کوواریانس به صورت زیر بدست می‌آیند:

$$\hat{R}_{XX}u(s) = \int_{T_1} \hat{r}_{XX}(s, t)u(t)dt, \quad u \in L_2(T_1),$$

$$\hat{R}_{XY}u(s) = \int_{T_Y} \hat{r}_{XY}(s, t)u(t)dt, \quad u \in L_2(T_Y), s \in T_X,$$

$$\hat{R}_{YY}u(s) = \int_{T_Y} \hat{r}_{YY}(s, t)u(t)dt, \quad u \in L_2(T_Y).$$

که در آن  $\hat{r}_{XX}(s, t)$  و  $\hat{r}_{YY}(s, t)$  به ترتیب توابع کوواریانس نمونه‌ای و  $\hat{r}_{XY}(s, t)$  تابع کوواریانس متقاطع نمونه است.

### ۳ تحلیل همبستگی کانونی هموار شده‌ی تابعی

همان‌طور که لرگانس و همکاران (۱۹۹۳) بیان نموده‌اند، هموارسازی در محاسبه‌ی همبستگی کانونی مفید بوده و ماکسیمم کردن  $\langle v, R_{YY}v \rangle$  با در نظر گرفتن همواری، اطلاعات مفیدی را در مورد داده‌ها یا مدل ارائه می‌کند. ابتدا تاوان همواری هر تابع را معرفی می‌کنیم.

تعریف. تاوان همواری هر تابع مانند  $f$  که به صورت  $[f, f]$  نشان داده می‌شود برابر است با

$$[f, f] = \int_I f^{(2)}(t)f^{(2)}(t)dt,$$

که در آن  $[,]$ ، حاصل ضرب داخلی روی فضای توابعی است که مشتق دوم آن‌ها موجود و متناهی است. همچنین  $f^{(2)}$  نشان‌دهنده‌ی مشتق مرتبه‌ی دوم تابع  $f$  است.

روش مستقیم برای به کارگیری همواری، تغییر شرایطی است که تحت آن ماکسیمم همبستگی بدست می‌آید. این کار با اضافه کردن عبارتی برای تاوان همواری به صورت زیر انجام می‌شود:

$$(۱) \quad \langle u, R_{XX}u \rangle + \alpha_1[u, u] = \langle v, R_{YY}v \rangle + \alpha_2[v, v] = ۱,$$

که در آن  $\alpha_1$  و  $\alpha_2$  پارامترهای همواری با مقدار مثبت هستند. این تغییر به گونه‌ای عمل می‌کند که در ارزیابی تغییرات کانونی نه تنها واریانس آن‌ها مورد توجه است بلکه مقدار همواری و مقایسه‌ی یک مجموع وزنی از این دو کمیت با عبارات کوواریانس نیز مورد توجه قرار می‌گیرد.

اکنون مسئله‌ی ماکسیمم سازی  $\langle u, R_{XX}u \rangle$  تحت قید (۱) معادل با ماکسیمم کردن همبستگی تاوانیده<sup>۳</sup> به صورت زیر خواهد بود:

$$\begin{aligned} \rho_k = & \sup \langle u, R_{YX}v \rangle = \langle u_k, R_{YX}v_k \rangle \\ & u \in L_2(T_X), \langle u, R_{XX}u \rangle + \alpha_1[u, u] = ۱, \\ & v \in L_2(T_Y), \langle v, R_{YY}v \rangle + \alpha_2[v, v] = ۱ \end{aligned}$$

از این راهکار به عنوان تحلیل همبستگی کانونی هموار شده یاد می شود. (SCCA)

انتخاب مناسب پارامترهای همواری به گونه ای است که توابع وزنی مرتبط با زوج متغیرهای کانونی نسبتاً هموار باشند و ضریب همبستگی به طور غیر منطقی کوچک نباشد. به طور واضح هر چه مقدار  $\alpha_1$  و  $\alpha_2$  بزرگتر انتخاب شود تأکید بیش تری روی تاوان همواری انجام می شود. از طرف دیگر، انتخاب مقادیر کوچک برای پارامترهای  $\alpha_1$  و  $\alpha_2$  موجب تأکید بیش تر بر ضریب همبستگی واقعی برای تغییرات خواهد شد. در این مقاله حالت خاص  $\alpha_1 = \alpha_2 = \alpha$  در نظر گرفته شده است که به سادگی محاسبات منجر می شود. برای انتخاب پارامتر همواری  $\alpha$  می توان از معیار اعتبارسنجی متقابل استفاده نمود. کافی است فرض کنید  $\hat{u}_\alpha^{(-i)}$  و  $\hat{v}_\alpha^{(-i)}$  مقدار ماکسیم کننده عبارت  $\widehat{\text{corr}}_\alpha^{-i}(u, v)$  است که در محاسبات آن، عنصر  $i$ ام نمونه حذف شده است. مقدار اعتبارسنجی متقابل برای  $\alpha$  به صورت توان دوم ضریب همبستگی  $n$  زوج  $(\hat{u}_\alpha^{(-1)T} X_1, \hat{v}_\alpha^{(-1)T} Y_1) \dots (\hat{u}_\alpha^{(-n)T} X_1, \hat{v}_\alpha^{(-n)T} Y_1)$  محاسبه می شود و  $\alpha$  طوری انتخاب می شود که مقدار این همبستگی ماکسیم شود.

مشابه بخش قبل، تحلیل همبستگی کانونی هموار شده، بر اساس اطلاعات نمونه به صورت زیر محاسبه می گردد:

$$\begin{aligned} \hat{\rho}_k &= \sup_{\substack{u \in L_r(T_1), \langle u, \hat{R}_{XX}u \rangle + \alpha[u, u] = 1, \\ v \in L_r(T_2), \langle v, \hat{R}_{YY}v \rangle + \alpha[v, v] = 1}} \langle u, \hat{R}_{YX}v \rangle \\ &= \langle u_k, \hat{R}_{YX}v_k \rangle. \end{aligned}$$

یا به طور معادل، تحلیل همبستگی کانونی هموار شده از ماکسیم کردن کسر زیر حاصل می شود:

$$(2) \quad \widehat{\text{corr}}_\alpha^2(u, v) = \frac{\langle u, \hat{R}_{XY}v \rangle^2}{(\langle u, \hat{R}_{XX}u \rangle + \alpha[u, u])(\langle v, \hat{R}_{YY}v \rangle + \alpha[v, v])}.$$

با فرض برقراری شرایط مرزی دوره ای روی مشتق دوم و سوم تابع دلخواه  $u$  و نیز شرط  $\int u^{(4)} dt < \infty$  داریم:  $\int u^{(2)} v^{(2)} dt = \int uv^{(4)} dt$  در نتیجه برای هر تابع دلخواه  $u$  که در این شرایط صدق کند می توان نوشت:

$$[u, u] = \langle u, D^* u \rangle,$$

که در آن  $D^*$  عملگر مشتق چهارم است.



توابع  $(\hat{u}_\alpha, \hat{v}_\alpha)$  که رابطه‌ی (۲) را ماکسیم می‌کند، با برقراری این شرایط، از طریق یافتن بزرگ‌ترین مقدار ویژه‌ی مثبت برای دستگاه معادلات زیر حاصل می‌شود:

$$(3) \quad \begin{pmatrix} \circ & \hat{R}_{XY} \\ \hat{R}_{YX} & \circ \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \rho \begin{pmatrix} \hat{R}_{XX} + \alpha D^* & \circ \\ \circ & \hat{R}_{YY} + \alpha D^* \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$$

برای حل دستگاه معادلات (۳) در عمل باید توابع  $u(t)$ ،  $v(t)$  و عملگرهای کوواریانس  $R(s, t)$  در یک شبکه مناسب گسسته‌سازی شود. در این مقاله ابتدا توابع و عملگر همواری با استفاده از سری‌های فوریه تقریب زده شده‌اند. همچنین عملگر  $D^*$  با تقریب تفاضل جایگزین شده است. در این صورت دستگاه (۳) به یک دستگاه خطی با ابعاد بزرگ تبدیل می‌شود که مقادیر و بردارهای ویژه متناظر با آن به‌وسیله‌ی روش‌های عددی استاندارد بدست می‌آیند.

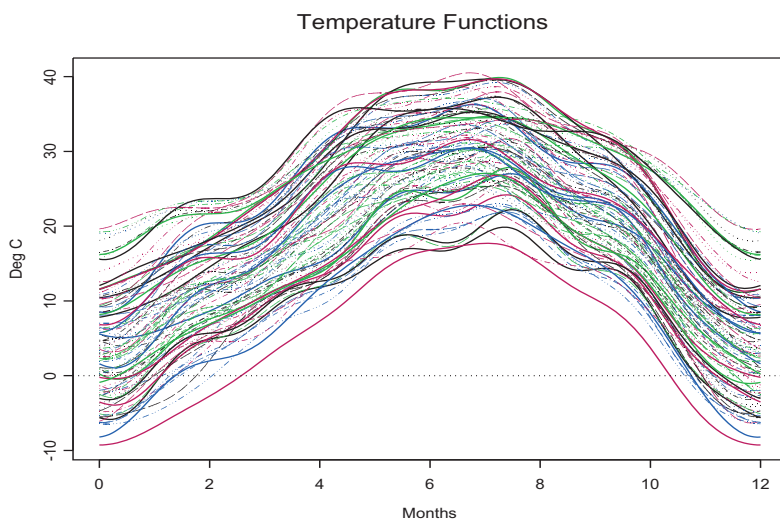
## ۴ تحلیل داده‌های واقعی

سه پدیده‌ی رطوبت، بارندگی و دما ماهیت تابعی دارند زیرا مقادیر رطوبت، بارندگی یا دما در هر منطقه به زمان بستگی دارد و در طول زمان تغییر می‌کند. از این رو استفاده از تحلیل داده‌های تابعی برای تحلیل آن‌ها توجیه‌پذیر است. داده‌هایی که در این مقاله مورد تحلیل قرار گرفته‌اند از داده‌هایی که سازمان هواشناسی ایران در سال ۲۰۰۶، از ۱۰۲ ایستگاه سینوپتیکی در نقاط مختلف کشور جمع‌آوری کرده است انتخاب شده‌اند. در این مجموعه داده، برای هر ایستگاه میانگین ماهانه‌ی رطوبت نسبی، دما و مقدار بارندگی در آن منطقه گزارش شده است (برای جزئیات بیش‌تر در مورد داده‌ها به تازیکه و حسینی نسب، ۱۳۸۷ رجوع کنید).

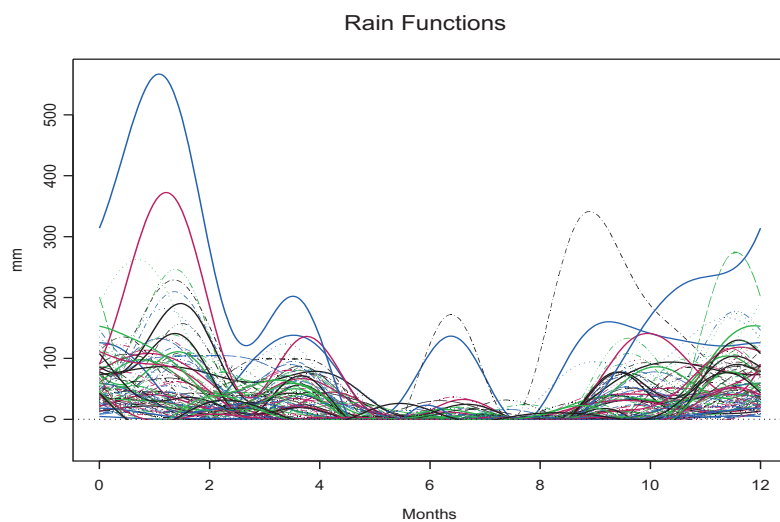
به دلیل ماهیت دوره‌ای (تناوبی) این سه پدیده، برای برازش منحنی به مجموعه داده‌ها از یک بسط سری فوریه با ۱۳ جمله استفاده نموده‌ایم. در تحلیل این داده‌ها، هدف بررسی و استخراج الگوهای مناسب همبستگی بین پدیده‌های رطوبت، دما و مقدار بارندگی است. این کار با استفاده از تحلیل همبستگی کانونی تابعی صورت می‌گیرد. نمودار دما، بارندگی و رطوبت برای این ۱۰۲ ایستگاه به‌ترتیب در شکل‌های ۱، ۲ و ۳ آورده شده است.

به‌منظور تحلیل اولیه‌ی داده‌ها می‌توان از نمودار تابع همبستگی و سطح تراز استفاده کرد. نمودار توابع همبستگی و سطوح تراز دما، رطوبت و بارندگی در شکل ۴ ترسیم شده است.

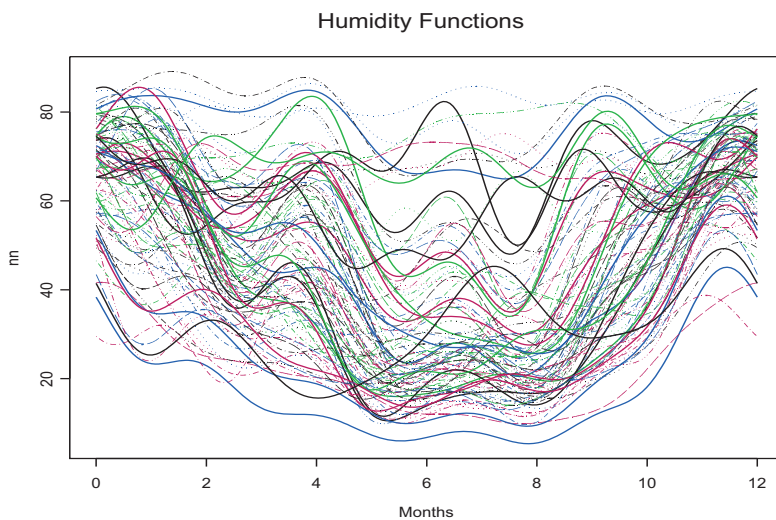
تابع همبستگی و نمودار سطوح تراز به‌منظور تفسیر مناسب‌تر در مقیاس هفتگی رسم شده‌اند. همان‌طور



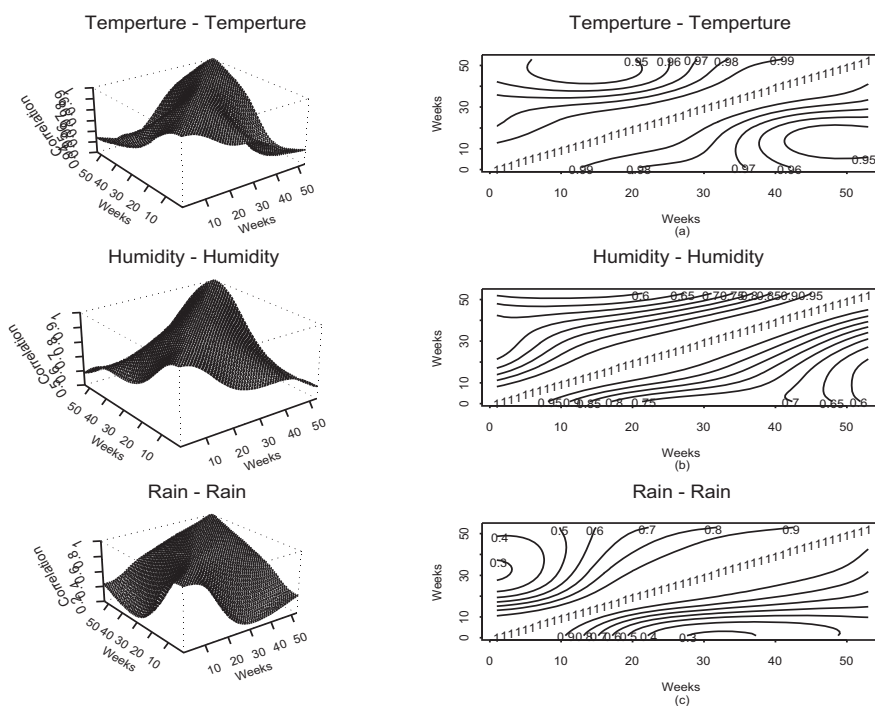
شکل ۱. نمودار دما برای ۱۰۲ ایستگاه هواشناسی در سال ۲۰۰۶



شکل ۲. نمودار بارندگی برای ۱۰۲ ایستگاه هواشناسی در سال ۲۰۰۶



شکل ۳. نمودار رطوبت برای ۱۰۲ ایستگاه هواشناسی در سال ۲۰۰۶



شکل ۴. نمودار تابع همبستگی و سطح تراز به ترتیب دما، رطوبت و بارندگی برای داده‌های ۱۰۲ ایستگاه هواشناسی در سال ۲۰۰۶

جدول ۱. جدول همبستگی کانونی دما و بارندگی

$\rho_6^*$	$\rho_5^*$	$\rho_4^*$	$\rho_3^*$	$\rho_2^*$	$\rho_1^*$
۰/۰۰۷	۰/۰۲۶	۰/۰۶۷	۰/۳۴۱	۰/۵۸۰	۰/۸۱۵

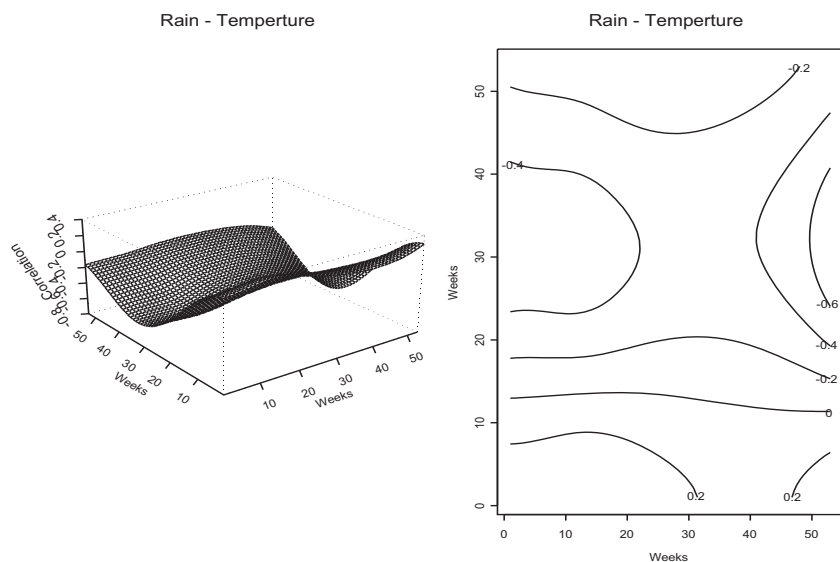
که نمودار (a) در شکل ۴ نشان می‌دهد با گذشت زمان از هفته‌های ابتدای سال میلادی (نیمه دوم دی‌ماه) به سمت هفته‌های انتهایی سال میلادی، همبستگی میان دما در این هفته‌ها با دما در چند هفته‌ی اول سال میلادی از ۰/۹۹ به ۰/۹۵ کاهش می‌یابد. نمودار (b) در شکل ۴ نشان می‌دهد با حرکت از هفته‌های ابتدایی سال میلادی به سمت هفته‌های پایانی آن، از میزان همبستگی میان رطوبت در این هفته‌ها و هفته‌های ابتدای سال کاسته می‌شود. همچنین نمودار (c) در شکل ۴ نشان می‌دهد با گذشت زمان از هفته ۱۰ سال میلادی به بعد مقدار همبستگی بارندگی در این زمان‌ها کم می‌شود به طوری که تا هفته‌ی سی‌ام همبستگی میان بارندگی در این هفته‌ها و هفته‌های ابتدایی سال میلادی از ۰/۹۰ به ۰/۳ کاهش می‌یابد. به‌علاوه چنان‌چه از هفته‌ی ۳۰ به طرفین حرکت کنیم به میزان همبستگی افزوده می‌شود.

در ادامه، با استفاده از تحلیل همبستگی کانونی ساختار همبستگی «دما و بارندگی»، «رطوبت و بارندگی» و «دما و رطوبت» به‌دقت مورد بررسی قرار می‌گیرد.

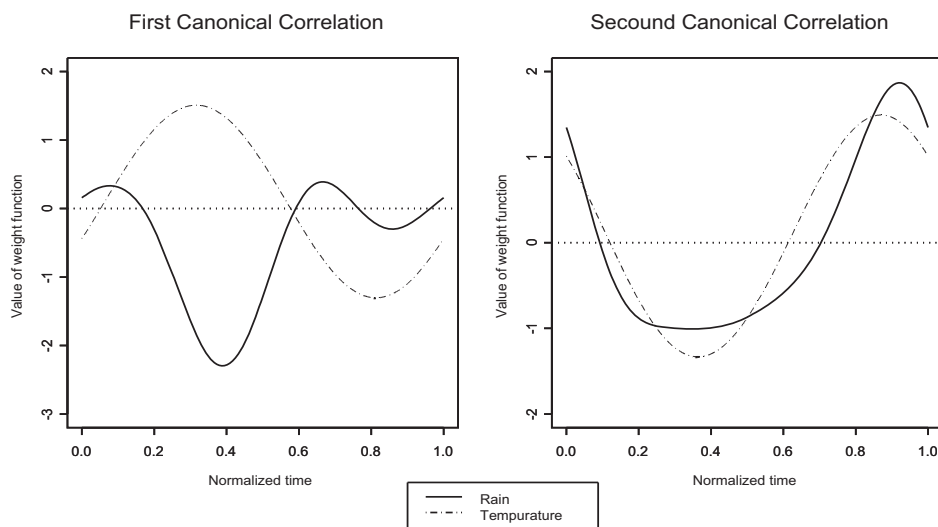
#### ۴/۱ تحلیل همبستگی کانونی دما و بارندگی

به‌منظور بررسی اولیه‌ی ساختار همبستگی داده‌های دما و بارندگی می‌توان از نمودارهای تابع همبستگی سه بُعدی و یا سطوح تراز آن استفاده کرد. شکل ۵ نمودار تابع همبستگی متقاطع و سطوح تراز دما و بارندگی را در ۱۰۲ ایستگاه نشان می‌دهد. محورهای افقی و عمودی در این نمودارها به‌ترتیب مقدار بارندگی و دما را بر حسب زمان (هفته) بیان می‌کند. همان‌طور که نمودار تراز نشان می‌دهد همبستگی میان دما و بارندگی اغلب عددی منفی است که تأییدی بر روابط واقعی بین این دو پدیده در طبیعت است. بر اساس شکل ۵، همبستگی میان دما و بارندگی در هفته‌های میانی سال میلادی (خرداد ماه) به بیش‌ترین مقدار خود رسیده است (۰/۶-) که به‌تدریج با فاصله گرفتن از هفته‌ی سی‌ام به طرفین از این میزان کاسته می‌شود.

برای تحلیل داده‌ها و استخراج الگوهای همبستگی بین دو پدیده‌ی دما و بارندگی از تحلیل همبستگی کانونی استفاده شده است. مقادیر همبستگی‌های کانونی این دو متغیر در جدول ۱ ارائه شده است. همان‌طور که ملاحظه می‌شود مقدار همبستگی‌های کانونی برای متغیرهای کانونی سوم و بعد از آن بسیار کوچک است لذا در ادامه‌ی تحلیل، تنها دو متغیر کانونی اول مورد استفاده قرار گرفته‌اند.



شکل ۵. نمودار تابع همبستگی و سطح تراز دما و بارندگی برای داده‌های ۱۰۲ ایستگاه هواشناسی در سال ۲۰۰۶



شکل ۶. نمودار توابع وزنی متغیرهای کانونی اول و دوم دما و بارندگی

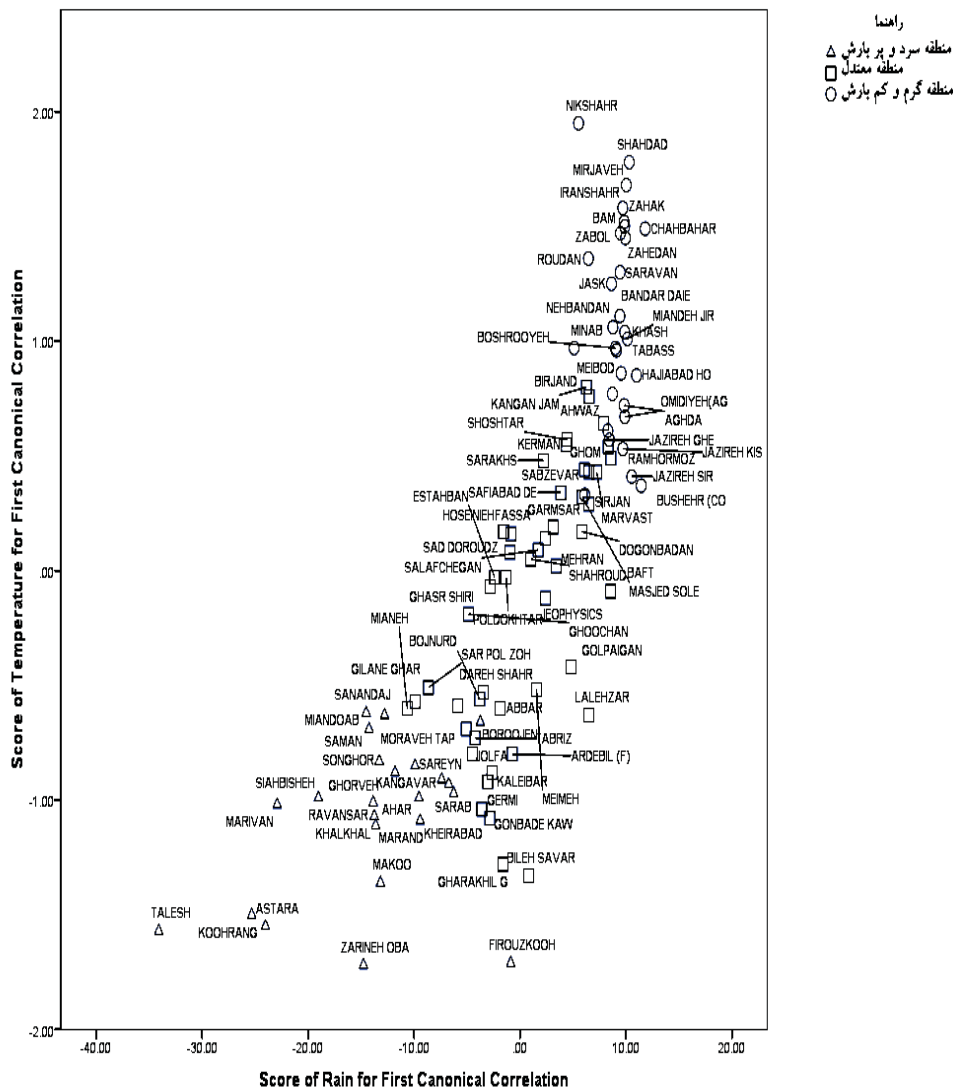
توابع وزنی متغیرهای کانونی اول و دوم در شکل ۶ رسم گردیده است. در این شکل زمان به گونه‌ای مقیاس‌بندی گردیده که این توابع مقادیر خود را در فاصله‌ی (۱, ۰) اختیار کنند.

نمودار سمت چپ از شکل ۶، توابع وزنی متغیرهای کانونی اول را نشان می‌دهد. همان‌طور که در این شکل نمایان است فرم توابع وزنی دو متغیر تا نقطه‌ی ۰/۷ (تا نیمه‌ی دوم ماه هشتم میلادی) متفاوت است به‌طوری که در ماه‌های ابتدایی سال میلادی، تابع وزنی بارندگی منفی و کاهش یافته ولی تابع وزنی دما مثبت و افزایش داشته است به‌نحوی که در پایان ماه پنجم میلادی بیش‌ترین اختلاف وزنی مشاهده می‌شود. به‌عبارت دیگر، این دو تابع را تا ماه هشتم میلادی می‌توان تقریباً قرینه در نظر گرفت. این همبستگی را می‌توان این‌گونه تفسیر نمود که ماه هشتم سال میلادی بارندگی و دما اثر معکوس دارند و این دو اثر صرف نظر از اولین تلاقی، در اوایل ماه هفتم میلادی (اوایل تیرماه) با هم تلاقی پیدا می‌کنند که نشان‌دهنده‌ی یکسان بودن اثر دما و بارندگی در همبستگی میان آن دو است. سپس تقریباً از میانه‌ی مرداد ماه به بعد اثرات این دو پدیده هم‌جهت می‌شود. بنا بر این می‌توان گفت از ابتدای سال میلادی (دی ماه) تا میانه‌ی مرداد ماه با افزایش دما، بارندگی کاهش و با کاهش دما، بارندگی افزایش می‌یابد و بالعکس.

با در نظر گرفتن این توابع وزن و به‌منظور تفسیر بهتر، نمودار پراکنش

$$(< u(t), X_i(t) >, < v(t), Y_i(t) >)$$

برای ۱۰۲ ایستگاه هواشناسی که به آن‌ها امتیاز<sup>۴</sup> ایستگاه‌ها گویند، برای اولین همبستگی کانونی در شکل ۷ رسم گردیده است. بنا بر آن‌چه گفته شد امتیاز هر ایستگاه دو مولفه دارد که مؤلفه‌ی اول آن بیان‌گر نقش و تأثیر اولین همبستگی کانونی بر بارندگی آن ایستگاه و مؤلفه‌ی دوم آن بیان‌کننده‌ی نقش و تأثیر اولین همبستگی کانونی بر دمای آن ایستگاه می‌باشد. نمودار پراکنش نشان می‌دهد ایستگاه نیکشهر دارای بیش‌ترین امتیاز دما و بارندگی است، بدین معنی که اولین همبستگی کانونی بر روی دما و بارندگی این ایستگاه نسبت به سایر ایستگاه‌ها، اثر بیش‌تری داشته است. این امر به‌علت بالا بودن دما در ماه‌های پنجم و ششم میلادی در این ایستگاه و بالا بودن میزان وزن دما در این ماه‌ها، همچنان‌که در نمودار سمت چپ شکل ۶ نمایان است، می‌باشد. این مطلب در مورد بارندگی نیز صحیح است. به‌علاوه ایستگاه تالش دارای کم‌ترین امتیاز دما و بارندگی است که نشان‌دهنده‌ی اثر بسیار کم اولین همبستگی کانونی بر روی دما و بارندگی این ایستگاه می‌باشد. این امر ناشی از بالا بودن میزان بارندگی در ماه نهم و دهم میلادی در این ایستگاه و پایین بودن میزان وزن بارندگی در این ماه‌ها، همچنان‌که در نمودار سمت چپ شکل ۶ نمایان است، می‌باشد. همچنین شکل ۷ نشان می‌دهد ۴۷ ایستگاه از ۱۰۲ ایستگاه دارای امتیاز دما



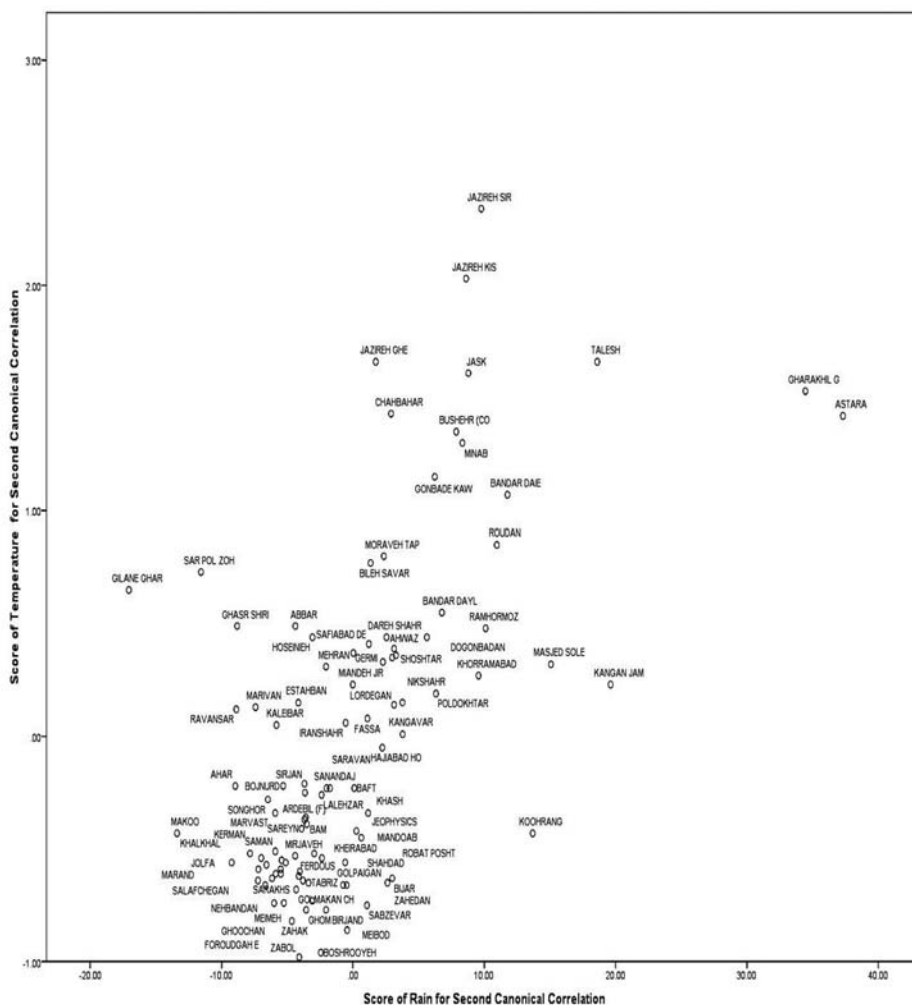
شکل ۷. نمودار امتیاز ۱۰۲ ایستگاه بر اساس اولین همبستگی کانونی دما و بارندگی

و بارندگی مثبت می‌باشند و امتیاز این ۴۷ ایستگاه به‌گونه‌ای است که با افزایش امتیاز بارندگی، امتیاز دما نیز افزایش می‌یابد. علاوه بر آن، با بررسی این نمودار و ویژگی‌های ایستگاه‌ها، نحوه‌ی قرارگیری ایستگاه‌ها در این نمودار را می‌توان ناشی از اقلیم آب و هوایی آن‌ها دانست. به این صورت که می‌توان ایستگاه را با توجه به تأثیر همبستگی کانونی اول بر روی دما و بارندگی آن‌ها در سه دسته جای داد: سرد و پربارش، معتدل و گرم و کم‌بارش، که این سه دسته در شکل نشان داده شده‌اند. یا به عبارت دیگر با حرکت از بالا سمت راست به پایین سمت چپ، روند امتیازها به‌گونه‌ای است که از میزان تأثیر همبستگی کانونی اول بر بارش ایستگاه‌ها کاسته و به میزان تأثیر همبستگی کانونی اول بر دمای ایستگاه‌ها افزوده می‌شود. بنا بر این اولین عامل همبستگی بین دما و بارندگی، متأثر از اقلیم آب و هوایی ایستگاه‌ها است.

نمودار سمت راست از شکل ۶، نمودار توابع وزنی متغیرهای کانونی دوم را نشان می‌دهد. همان‌طور که در شکل مشخص است در ماه‌های میانی سال میلادی، بارندگی زودتر از خط صفر عبور کرده است یا به عبارت دیگر ابتدا بارندگی و به دنبال آن دما از صفر عبور می‌کند. از این موضوع می‌توان چنین استنتاج نمود که در ماه‌های میانی سال میلادی تغییرات دما ناشی از تغییرات بارندگی است. همچنین شکل ۶ نشان می‌دهد که برخلاف تابع وزنی متغیر کانونی اول، روند تغییرات توابع وزنی دما و بارندگی یکسان است تا جایی که کم‌ترین و بیش‌ترین میزان وزن‌ها نیز در یک مقطع زمانی برای دما و بارندگی رخ داده است.

امتیازهای دومین متغیرهای کانونی دما و بارندگی در شکل ۸ ترسیم گردیده است. این شکل نشان می‌دهد تقریباً ۵۰ ایستگاه از ۱۰۲ دارای امتیازهای منفی هستند. اما تراکم ایستگاه‌های دارای امتیاز مثبت بیش‌تر از ایستگاه‌های دارای امتیازهای منفی است. به این صورت که امتیازهای منفی مقادیر خود را در فاصله‌ی  $(-25, 0)$  برای بارندگی و  $(-1, 0)$  برای دما اختیار می‌کنند درحالی که امتیازهای مثبت مقادیر خود را در فاصله‌ی  $(0, 40)$  برای بارندگی و  $(0, 2/5)$  برای دما اختیار کرده‌اند. این امر نشان‌دهنده‌ی اثر تقریباً مشابه دما و بارندگی در ایستگاه‌های دارای امتیاز منفی نسبت به ایستگاه‌های دارای امتیاز مثبت است. همان‌طور که در شکل نمایان است ایستگاه زابل دارای کم‌ترین امتیاز دما است. این امر ناشی از بالا بودن دما در ماه پنجم میلادی این ایستگاه نسبت سایر ماه‌های سال و پایین بودن تابع وزن دما در این ماه می‌باشد (نمودار سمت راست شکل ۶ را ببینید). همان‌طور که در شکل ۸ مشاهده می‌گردد ایستگاه آستارا دارای بیش‌ترین امتیاز بارندگی است. این امر ناشی از بالا بودن میزان بارندگی در ماه نهم میلادی این ایستگاه و بالا بودن میزان وزن بارندگی در این ماه است (نمودار سمت راست شکل ۶ را ببینید). همان‌طور





شکل ۸. نمودار امتیاز ۱۰۲ ایستگاه بر اساس دومین همبستگی کانونی دما و بارندگی

جدول ۲. جدول همبستگی کانونی رطوبت و بارندگی

$\rho_6^*$	$\rho_5^*$	$\rho_4^*$	$\rho_3^*$	$\rho_2^*$	$\rho_1^*$
۰/۰۱۸	۰/۰۹۳	۰/۱۸۶	۰/۳۵۶	۰/۶۰۴	۰/۶۹۶

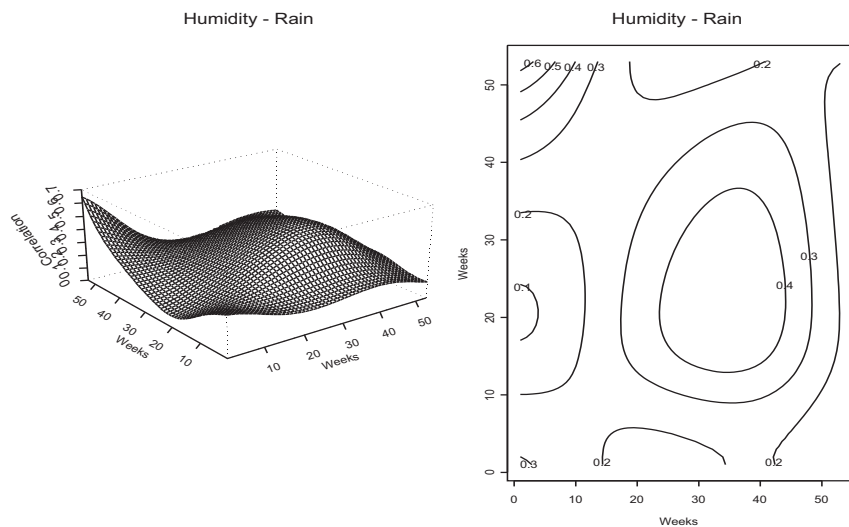
که می‌دانیم عوامل متعددی از قبیل نزدیکی به دریا، ارتفاع، شیب، ابری بودن هوا و ... در میزان تأثیر دما و رطوبت مؤثر می‌باشند. هر چند که امتیازها در شکل ۸ متأثر از تک تک این عوامل نمی‌باشد اما تأثیر همه‌ی این عوامل با هم چینش امتیازهای ایستگاه‌ها به‌گونه‌ای که در این شکل شاهد هستیم را باعث شده است. بدین معنی که ایستگاه‌هایی که تقریباً دارای امتیاز یکسان دما و بارندگی هستند برخی دارای ارتفاع نزدیک به هم هستند، برخی فاصله‌ی یکسان تا دریا دارند و ... در اصطلاح هواشناسی، تأثیر توأم عوامل ذکر شده را توپوگرافی می‌نامند. در نتیجه نحوه‌ی قرارگیری ایستگاه‌ها بر اساس امتیازها در شکل ۸ متأثر از توپوگرافی ایستگاه‌ها است.

## ۴/۲ تحلیل همبستگی کانونی رطوبت و بارندگی

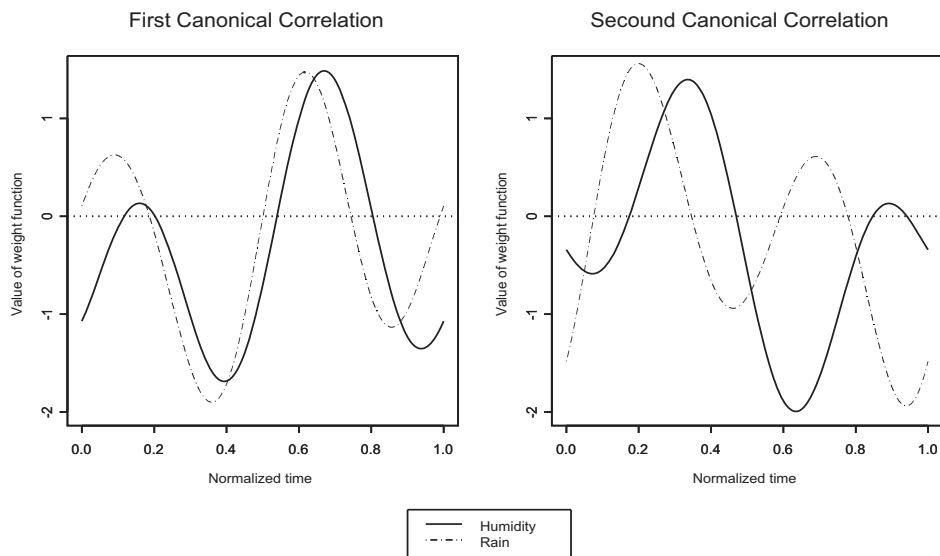
برای درک اولیه‌ی ساختار همبستگی داده‌های رطوبت و بارندگی، ابتدا نمودارهای تابع همبستگی سه بُعدی یا سطوح تراز آن را مطالعه می‌کنیم. شکل ۹ نمودار تراز همبستگی رطوبت و بارندگی را در ۱۰۲ ایستگاه نشان می‌دهد که در آن محورهای افقی و عمودی به‌ترتیب نشان‌دهنده‌ی رطوبت و بارندگی می‌باشد. این شکل نشان می‌دهد با گذشت زمان از هفته‌های ابتدایی سال میلادی حد اکثر تا هفته‌ی ۱۵ آن، همبستگی میان رطوبت در این هفته‌ها و بارندگی در هفته‌های پایانی سال میلادی از ۰/۶ به ۰/۳ کاهش می‌یابد.

با انجام تحلیل همبستگی کانونی برای تحلیل داده‌ها و استخراج الگوهای همبستگی بین دو پدیده‌ی رطوبت و بارندگی، مقادیر همبستگی کانونی بدست آمده‌اند که در جدول ۲ قابل مشاهده می‌باشند. به‌دلیل آن که مقدار همبستگی‌های کانونی برای متغیرهای کانونی سوم و بعد از آن نسبت به دو همبستگی کانونی اول کوچک است تنها دو متغیر کانونی اول در تحلیل مورد استفاده قرار گرفته‌اند. پس از محاسبه‌ی مقادیر همبستگی کانونی، توابع وزنی متغیرهای کانونی اول و دوم برای داده‌های رطوبت و بارندگی در شکل ۱۰ رسم شده است.

نمودار سمت چپ از شکل ۱۰، توابع وزنی متغیرهای کانونی اول را نشان می‌دهد. همان‌طور که در این شکل نمایان است نوسانات تابع وزنی رطوبت و بارندگی تقریباً یکسان است. این شکل نشان می‌دهد



شکل ۹. نمودار تابع همبستگی و سطح تراز رطوبت و بارندگی برای داده‌های ۱۰۲ ایستگاه هواشناسی در سال ۲۰۰۶



شکل ۱۰. نمودار توابع وزنی متغیرهای کانونی اول و دوم داده‌های رطوبت و بارندگی

کمترین مقدار تابع وزنی بارندگی کمی زودتر از کمترین مقدار رطوبت اتفاق می‌افتد. این مطلب در مورد بیشترین مقدار نیز صادق است و بیانگر آن است که تغییرات بارندگی به دنبال تغییرات رطوبت اتفاق می‌افتد. به عبارت دیگر تغییرات رطوبت ناشی از تغییرات بارندگی است.

نمودار پراکنش امتیازهای رطوبت در مقابل امتیازهای بارندگی ۱۰۲ ایستگاه هواشناسی برای همبستگی کانونی اول شکل ۱۱ رسم گردیده است. بر اساس این نمودار، تقریباً ۶۰ ایستگاه دارای امتیاز رطوبت و بارندگی مثبت هستند که نشان‌دهنده‌ی اثر بالای همبستگی کانونی اول بر رطوبت و بارندگی این ایستگاه‌ها است. همان‌طور که در شکل ۱۱ نمایان است با افزایش امتیاز رطوبت، امتیاز بارندگی افزایش می‌یابد. شکل نشان می‌دهد ایستگاه کنگاور دارای کمترین امتیاز رطوبت است. علت این امر پایین بودن میزان رطوبت در ماه‌های هفتم و هشتم میلادی در این ایستگاه و بالا بودن میزان وزن رطوبت در این ماه (مطابق با نمودار سمت چپ شکل ۱۰) است. با بررسی این نمودار و ویژگی‌های ایستگاه‌ها، نحوه‌ی قرارگیری امتیاز ایستگاه‌ها در این نمودار را می‌توان ناشی از اقلیم آب و هوایی آن‌ها دانست، به این صورت که می‌توان ایستگاه را در سه دسته جای داد: پربارش، بارش متوسط و کم‌بارش، که این سه دسته در شکل نشان داده شده‌اند. به عبارت دیگر، با حرکت از بالا سمت راست به پایین سمت چپ روند امتیاز ایستگاه‌ها به گونه‌ای خواهد بود که بر میزان بارش ایستگاه‌ها افزوده می‌شود. بنا بر این اولین عامل همبستگی میان بارندگی و رطوبت متأثر از اقلیم آب و هوایی ایستگاه‌ها است.

نمودار سمت راست از شکل ۱۰، توابع وزنی متغیرهای کانونی دوم را نشان می‌دهد. روند تغییرات این دو تابع تقریباً یکسان می‌باشد و می‌توان با یک تبدیل مناسب، این دو تابع را بر هم منطبق کرد. بنا بر این می‌توان گفت که در همبستگی دوم نیز، تغییرات رطوبت ناشی از تغییرات بارندگی است.

نمودار پراکنش امتیازهای دومین متغیرهای کانونی رطوبت و بارندگی در شکل ۱۲ ترسیم شده است. این شکل نشان می‌دهد که تقریباً ۸۵ ایستگاه امتیازهای خود را در فاصله‌ی (۱۰، ۵-) برای رطوبت و (۱۸، ۱۰-) برای بارندگی اختیار می‌کند. شکل نشان می‌دهد ایستگاه تالش دارای کمترین امتیاز بارندگی است که این امر به علت پایین بودن میزان بارندگی در ماه هشتم میلادی و بالا بودن وزن بارندگی در این ماه (مطابق با نمودار سمت راست شکل ۱۰) می‌باشد. با توجه به ویژگی‌های ایستگاه‌ها می‌توان گفت با حرکت از بالا سمت چپ به پایین سمت راست، ایستگاه‌ها به گونه‌ای قرار گرفته‌اند که از دریا فاصله گرفته و مرتفع‌تر می‌شوند. در نتیجه عامل همبستگی دوم میان رطوبت و بارندگی متأثر از میزان نزدیکی ایستگاه به دریا و ارتفاع ایستگاه می‌باشد.





جدول ۳. جدول همبستگی کانونی دما و رطوبت

$\rho_1^*$	$\rho_2^*$	$\rho_3^*$	$\rho_4^*$	$\rho_5^*$	$\rho_6^*$
۰/۸۷۸	۰/۸۱۴	۰/۴۸۶	۰/۰۵۵	۰/۰۱۱	۰/۰۰۲

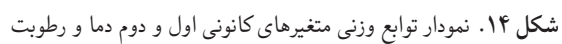
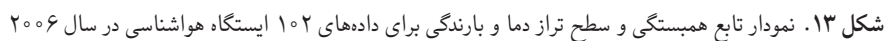
### ۴/۳ تحلیل همبستگی کانونی دما و رطوبت

بررسی نمودار تابع همبستگی سه بُعدی یا سطوح تراز داده‌های دما و رطوبت به فهم بهتر ساختار همبستگی آن‌ها کمک می‌کند. شکل ۱۳ نمودار تراز همبستگی میان دما و رطوبت را برای ۱۰۲ ایستگاه هواشناسی نشان می‌دهد که در آن محورهای افقی و عمودی به ترتیب نشان‌دهنده رطوبت و دما می‌باشد. همان‌طور که در این شکل نمایان است همبستگی میان دما و رطوبت در اغلب موارد، عددی منفی است که تأیید کننده روابط این دو پدیده در جهان واقعی است. همچنین این شکل نشان می‌دهد با حرکت از هفته‌ی ۲۵ (خرداد ماه) به طرفین، از همبستگی میان دما در این هفته‌ها و رطوبت در هفته‌های ابتدایی سال میلادی (اواسط دی ماه) کاسته می‌شود به‌طوری که در هفته ۴۵ (اوایل آبان ماه) این همبستگی به صفر می‌رسد. این بدان معنی است که پیش‌بینی دما در هفته‌ی ۴۵ سال میلادی با استفاده از رطوبت در هفته‌ی ابتدای سال بسیار مشکل است.

استخراج الگوهای همبستگی بین دو پدیده رطوبت و بارندگی بر اساس تحلیل همبستگی کانونی انجام شده است و مقادیر همبستگی کانونی بدست آمده‌اند که در جدول ۳ قابل مشاهده می‌باشند. به دلیل آن که مقدار همبستگی‌های کانونی برای متغیرهای کانونی سوم و بعد از آن نسبت به دو همبستگی کانونی اول کوچک است در تحلیل تنها دو متغیر کانونی اول مورد استفاده قرار گرفته‌اند.

پس از محاسبه‌ی میزان همبستگی کانونی، توابع وزنی متغیرهای کانونی اول و دوم برای داده‌های دما و رطوبت در شکل ۱۴ رسم شده‌اند.

نمودار سمت چپ از شکل ۱۴، توابع وزنی متغیرهای کانونی اول را نشان می‌دهد. بر اساس این شکل، توابع وزنی دما و رطوبت نسبت به هم قرینه‌اند، به این معنی که بیش‌ترین میزان دما دقیقاً در زمانی اتفاق می‌افتد که تابع وزنی رطوبت کم‌ترین مقدار خود را اختیار می‌کند و با افزایش تابع وزنی دما تابع وزنی رطوبت به همان میزان کاهش می‌یابد. این امر بیانگر اثر مخالف این دو نسبت به یکدیگر است. با بررسی نمودار امتیازهای اولین متغیرهای کانونی دما و رطوبت می‌توان گفت تقریباً ۴۵ ایستگاه





دارای امتیازهای دما و بارندگی مثبت هستند و با افزایش امتیاز رطوبت، امتیاز دما افزایش یافته است. همچنین ایستگاه چابهار دارای کمترین امتیاز رطوبت است که این امر به علت پایین بودن میزان رطوبت این ایستگاه در ماه‌های ابتدا و انتهای سال و بالا بودن تابع وزن رطوبت در ابتدا و انتهای نمودار سمت چپ از شکل ۱۴ می‌باشد. با توجه به بررسی این امتیازها و ویژگی‌های ایستگاه‌ها، نحوه‌ی قرارگیری ایستگاه‌ها در این نمودار را باز هم می‌توان ناشی از اقلیم آب و هوایی آن‌ها دانست، به این صورت که می‌توان ایستگاه را با توجه به تأثیرشان در همبستگی میان رطوبت و دما در سه دسته جای داد: گرم و خشک، معتدل و سرد و پربارش. به عبارت دیگر، با حرکت از بالا سمت راست به پایین سمت چپ، روند ایستگاه‌ها به گونه‌ای خواهد بود که از میزان دما و رطوبت ایستگاه‌ها کاسته می‌شود.

نمودار سمت راست از شکل ۱۴، توابع وزنی متغیرهای کانونی دوم را نشان می‌دهد. این دو تابع وزنی تقریباً به نوعی قرینه‌اند و تنها در ابتدای سال میلادی روند تغییرات دو تابع متفاوت از سایر ماه‌های سال می‌باشد.

با بررسی نمودار امتیازهای دومین متغیرهای کانونی دما و رطوبت می‌توان گفت با حرکت از سمت چپ نمودار به سمت راست با افزایش دما و افزایش رطوبت روبه‌رو هستیم. امتیازهای بدست آمده نشان می‌دهند، ایستگاه زرینه دارای کمترین امتیاز دما است که این امر به علت پایین بودن تابع وزن دما در فاصله‌ی بین ماه‌های ۶ تا ۹ میلادی و بالا بودن میزان دمای این ایستگاه در این ماه‌ها است (نمودار سمت راست شکل ۱۴ را ببینید). نحوه قرارگیری ایستگاه‌ها در نمودار امتیازات متأثر از توپوگرافی آن‌ها بوده است اما تفاوت آن با نمودار امتیازات دومین همبستگی کانونی میان دما و رطوبت در آن است که در این نمودار ویژگی‌های ایستگاه‌ها به صورت دیگری در چیدمان ایستگاه‌ها تأثیرگذار بوده‌اند، به عنوان مثال، عامل ساعت‌های تابش آفتاب و تعداد روزهایی با آسمان صاف باعث نزدیکی امتیازهای دومین همبستگی کانونی میان دما و بارندگی برخی ایستگاه‌ها گردیده که همین عامل باعث شده است تا امتیازهای دومین همبستگی کانونی میان رطوبت و دما برای برخی دیگر از ایستگاه‌ها به هم نزدیک شوند. بنا بر این ویژگی‌های ایستگاه‌ها در هر عامل همبستگی، نمودهای مختلفی ایجاد کرده است.

## ۵ نتیجه‌گیری

با توجه به بررسی ساختار همبستگی میان دما، رطوبت و بارندگی مربوط به ۱۰۲ ایستگاه هواشناسی ایران در سال ۲۰۰۶، از طریق تحلیل همبستگی کانونی تابعی می‌توان گفت: در بررسی رابطه‌ی میان دما و بارندگی اولین عامل همبستگی، ناشی از اقلیم آب و هوایی ایستگاه‌ها و دومین عامل همبستگی ناشی

از توپوگرافی آن‌ها می‌باشد. در بررسی رابطه‌ی میان رطوبت و بارندگی، اولین عامل متأثر از اقلیم آب و هوایی (بارش) و دومین عامل متأثر از ارتفاع و نزدیکی به دریا ایستگاه‌ها می‌باشد. همچنین در بررسی رابطه‌ی میان دما و رطوبت، اولین عامل همبستگی مربوط به اقلیم آب و هوایی (بارش سالانه و دمای سالانه) و دومین عامل همبستگی با توپوگرافی ایستگاه‌ها مرتبط است.

## توضیحات

۱. Cross-covariance function

۲. Cross-correlation function

۳. Penalized Correlation

۴. Score

## مرجع‌ها

تازیکه، ن.، خیراله زاده، ن.، موقری، د. و حسینی نسب، م. (۱۳۸۷). تحلیل داده‌های تابعی: تعمیمی از تحلیل داده‌های چندمتغیره، *مجله اندیشه‌ی آماری*، ۲۵، ۵۲-۶۸.

Brillinger, D.R. (1975). *Time Seris, Data Analysis and Theory*. Hlot, Rinehart and Winston, New York.

Dauxois, J. and Nkiet, G.M. (1997). Canonical analysis of two Euclidean subspaces and its applications. *Linear Algebra Appl.*, **264**, 355-388.

He, G., Müller, H.G. and Wang, J.L. (2003). Functional canonical analysis for square integrable stochastic processes. *Multivariate Analysis.*, **85**, 54-77.

Hannan, E.J. (1961). The general theory of canonical correlation and its relation to functional analysis. *J. Austral. Math. Soc.*, **2**, 229-242.

Hottelling, H. (1936). Relation between two sets of variates. *Biometrika*, **28**, 321-377.

Leurgans, S.E. and Moyeed, R.A. and Silverman, B.W. (1993). Canonical corelation analysis when data are curves. *J. R. statist. Soc. B*, **55**, 725-740.

Ramsay, J.O. and Silverman, B.W. (2005). *Functional data analysis*. Springer, 2nd edition, New York.

Ramsay, J.O. (1982). When data are functions. *Psychometrika*, **47**, 379-396.

Rice, J.A. and Silverman, B.W. (1991). Estimating the mean and covariance structure non-parametrically when the data are curves. *J. Roy. statist. Soc. B*, **53**, 233-243.

Service, S.K., Rice, J.A. and Ghavez, F.P. (1998). Relation between physical and biological variables during the upwelling period in Monterey Bay. *Ca, Deep-Sea Research II Vol.*, **45**, 1669-1685.

سید محمد ابراهیم حسینی نسب  
گروه آمار، دانشکده‌ی علوم ریاضی،  
دانشگاه شهید بهشتی،  
تهران، ایران.

رایانشانی: [m\\_hosseininasab@sbu.ac.ir](mailto:m_hosseininasab@sbu.ac.ir)

آناهیتا متقی گلشن  
گروه آمار، دانشکده‌ی علوم ریاضی،  
دانشگاه شهید بهشتی،  
تهران، ایران.

محمد رضا فرید روحانی  
گروه آمار، دانشکده‌ی علوم ریاضی،  
دانشگاه شهید بهشتی،  
تهران، ایران.