

معادلات پیش‌بینی برای شبکه‌های عصبی خود‌بازگشت با استفاده از فیلتر کالمن تعمیم یافته

محمد رضا یگانگی، رحیم چینی پرداز
گروه آمار، دانشگاه شهید چمران اهواز

بسیاری از مدل‌های خطی و غیر خطی سری زمانی و همچنین سیستم‌های پویا را می‌توان به صورت یک مدل فضای حالت نمایش داد و از ابزارهای موجود برای مدل‌های فضای حالت استفاده کرد. از جمله می‌توان از نمایش فضای حالت برای مدل‌های سری زمانی مبتنی بر شبکه‌های عصبی یا شبکه‌های عصبی خودبازگشت استفاده کرد. با استفاده از همین ویژگی روش‌های مختلفی مبتنی بر نمایش فضای حالت و فیلتر کالمن برای برآورد ضرایب چنین مدل‌هایی در منابع ارائه شده است. با این وجود به مسئله پیش‌بینی در این مدل‌ها با استفاده از فیلتر کالمن کمتر توجه شده. در این مقاله شبکه‌های عصبی خودبازگشت در قالب مدل‌های فضای حالت مورد بررسی قرار گرفته و با استفاده از فیلتر کالمن تعمیم یافته روش جدیدی برای پیش‌بینی در این مدل‌ها ارائه خواهد شد.

واژه‌های کلیدی: سری‌های زمانی، شبکه‌های عصبی، مدل فضای حالت، فیلتر کالمن تعمیم یافته، پیش‌بینی.

۱ مقدمه

مدل‌های سری زمانی خطی، به دلیل سادگی، بسیار مورد توجه بوده‌اند. در سال‌های اخیر به دلیل ضعف این مدل‌ها در توجیه و پیش‌بینی فرآیندهای غیر خطی، گرایش‌هایی به سمت یافتن مدل‌های غیر خطی به وجود آمد. از اواخر دهه ۸۰، مدل‌های شبکه عصبی به عنوان یکی از روش‌های رگرسیون غیر خطی، برای مدل‌سازی و پیش‌بینی سری‌های زمانی غیر خطی مورد توجه قرار گرفتند. این مدل‌ها به دلیل توانایی در تقریب زدن بسیاری از توابع غیر خطی و همچنین سهولت در استفاده، به طور وسیعی در مدل‌سازی سری‌های زمانی اقتصادی کاربرد پیدا کرده‌اند (فرانسس و ون‌دیک ۲۰۰۰).

مدل‌های فضای حالت اولین بار توسط کالمن^۱ (۱۹۶۰) و کالمن و بوکسی (۱۹۶۱) معرفی شدند. این مدل‌ها در حقیقت یک ساختار کلی برای فرمول‌بندی سیستم‌های پویای خطی یا غیر خطی هستند. استفاده از نمایش فضای حالت این امکان را ایجاد می‌کند که خطاهای اندازه‌گیری نیز وارد مدل شوند. همچنین با استفاده از روش‌هایی مثل بیشینه درست‌نمایی و تجزیه خطای پیش‌بینی، می‌توان معادلات پیش‌بینی را تشکیل داد. به علاوه در بسیاری از مدل‌های غیر خطی، ضرایب مدل ممکن است در طول زمان تغییر کنند. به این ترتیب با استفاده از مدل فضای حالت، با اضافه شدن مشاهدات جدید علاوه بر اصلاح پیش‌بینی‌ها، می‌توان ضرایب مدل را نیز بر اساس میزان خطای پیش‌بینی و مشاهدات جدید به روز رسانی کرد (پریستلی ۱۹۸۸). به دلیل این ویژگی‌ها، استفاده از نمایش فضای حالت و فیلتر کالمن برای پیش‌بینی در سری‌های زمانی غیرخطی و ناپایدار می‌تواند به افزایش دقت در انجام پیش‌بینی و افزایش قدرت انطباق‌پذیری مدل کمک کند. علاوه بر این، استفاده از فضای حالت و فیلتر کالمن به تعمیم مدل‌های غیر خطی و افزایش کارایی الگوریتم‌های آموزش نیز کمک خواهد کرد.

در دو دهه اخیر نمایش فضای حالت و فیلتر کالمن در شبکه‌های عصبی، به عنوان روشی برای برآورد ضرایب شبکه، بسیار مورد توجه قرار گرفته است. سینگال و وو (۱۹۸۹) و پاسکوریوس و فلدکمپ (۱۹۹۱ و ۱۹۹۷) با استفاده از نمایش فضای حالت، فیلتر کالمن تعمیم یافته و فیلتر کالمن تعمیم یافته مجزا الگوریتم‌هایی را برای آموزش شبکه‌های عصبی پیش‌خور چند لایه ارائه کرده‌اند. پلمر (۱۹۹۵) و سان و مارکو (۱۹۹۸) نیز از نسخه‌های دیگر فیلتر کالمن برای آموزش شبکه‌های عصبی استفاده کرده‌اند. نکته قابل توجه در مورد این روش‌ها، قدرت انطباق با مشاهدات جدید و برآورد پارامترهای مدل به صورت برخط است (پاسکوریوس و فلدکمپ ۲۰۰۱).

استفاده از مدل‌های فضای حالت و فیلتر کالمن برای پیش‌بینی در مدل‌های شبکه عصبی کمتر مورد توجه قرار گرفته است. در این مقاله مدل‌های سری زمانی مبتنی بر شبکه‌های عصبی در قالب مدل‌های فضای حالت مورد بررسی قرار گرفته و با استفاده از فیلتر کالمن تعمیم یافته معادلات پیش‌بینی و به روز رسانی برای این مدل‌ها ارائه خواهد شد. در ادامه مقاله، در بخش دوم فرمول‌بندی مدل‌های فضای حالت و فیلتر کالمن و فیلتر کالمن تعمیم یافته ارائه می‌شود. بخش سوم به بررسی شبکه‌های عصبی در قالب مدل‌های فضای حالت و فرمول‌بندی معادلات جدید پیش‌بینی و به روز رسانی مبتنی بر فیلتر کالمن تعمیم یافته، برای پیش‌بینی در شبکه‌های عصبی می‌پردازد. برای بررسی کارایی روش ارائه

¹Kalman

شده، با استفاده از معادلات بخش سوم، در بخش چهارم به پیش‌بینی سری زمانی لکه‌های خورشیدی پرداخته می‌شود. یک جمع بندی کلی و پیشنهاداتی برای استفاده از مدل‌های فضای حالت در بخش پنجم ارائه شده است.

۲ مدل‌های فضای حالت و فیلتر کالمن

در یک مدل فضای حالت، فرض اصلی آن است که سری زمانی مورد بررسی، X_t ، تحت تأثیر یک فرآیند اغتشاش محض و یک فرآیند مارکف مرتبه اول تغییر می‌کند. یعنی یک مدل فضای حالت را می‌توان با دو معادله به صورت زیر نشان داد.

$$\alpha_t = f(t, \alpha_{t-1}) + v_t \quad (۱)$$

$$X_t = h(t, \alpha_t) + w_t \quad (۲)$$

معادله (۱) را معادله فرآیند یا معادله حالت و معادله (۲) را معادله اندازه می‌نامند. در این مدل w_t و v_t دو فرآیند مستقل اغتشاش محض گوسی با میانگین صفر و ماتریس‌های کوواریانس به ترتیب R_t و Q_t هستند یعنی $Cov(v_t, w_t) = 0$. $\{\alpha_t\}$ یک فرآیند مارکف مرتبه اول M متغیری و X_t یک سری زمانی N متغیری است. فرم تابعی $f(t, \alpha_{t-1})$ در حالت کلی ماتریس انتقال غیر خطی است که در طول زمان تغییر می‌کند و $h(t, \alpha_t)$ در حالت کلی ماتریس غیر خطی اندازه نامیده می‌شود که ممکن است در طول زمان تغییر کند. در حالت خاصی که ماتریس‌های انتقال و اندازه خطی باشند، مدل فضای حالت به صورت زیر است:

$$\alpha_t = F_{t-1,t} \alpha_{t-1} + v_t \quad (۳)$$

$$X_t = H_t \alpha_t + w_t \quad (۴)$$

که در آن $F_{t-1,t}$ یک ماتریس انتقال و H_t یک ماتریس اندازه است. مدل فضای حالتی که با روابط (۳) و (۴) نمایش داده می‌شود کاملاً با استفاده از عملگرهای خطی تعریف شده ولی بسیاری از حالت‌های غیر خطی را نیز شامل می‌شود.

۱-۲ فیلتر کالمن

در استفاده از فیلتر کالمن، هدف حل همزمان معادلات فرآیند و اندازه (۳) و (۴) برای یافتن حالت بهینه α است. به بیان دیگر، می‌خواهیم بر اساس داده‌های مشاهده شده y_1, y_2, \dots, y_k ، برآورد کمترین مربعات خطای حالت فرآیند مارکف در لحظه i ، یعنی α_i ، را بدست آوریم. مسئله فیلتر کردن نامیده می‌شود اگر $i = k$ ، پیش‌بینی نامیده می‌شود اگر $i > k$ و هموارسازی نامیده می‌شود اگر $1 \leq i < k$ باشد.

فرض کنید که سری زمانی X_t که با معادلات (۳) و (۴) نمایش داده شد، در زمان $t-1$ قرار داشته باشد و بخواهیم وضعیت سری زمانی و حالت فرآیند مارکف را برای لحظه t پیش‌بینی کنیم. در این صورت معادلات پیش‌بینی به صورت زیر قابل فرمول‌بندی هستند:

$$\hat{\alpha}_{t|t-1} = F_{t-1,t} \hat{\alpha}_{t-1|t-1} \quad (5)$$

$$P_{t|t-1} = F_{t-1,t} P_{t-1|t-1} F'_{t-1,t} + R_{t-1} \quad (6)$$

$$\hat{X}_t = H_t \hat{\alpha}_{t|t-1} \quad (7)$$

که در آنها $\hat{\alpha}_{i|j}$ ، برآورد حالت فرآیند مارکف در لحظه i بر اساس مشاهدات X_1, X_2, \dots, X_j است. همچنین $P_{i|j}$ ، ماتریس کوواریانس بردار خطای حالت، $\tilde{\alpha}_t = \alpha_t - \hat{\alpha}_t$ ، در زمان i بر اساس مشاهدات X_1, X_2, \dots, X_j است:

$$P_{i|j} = E\{(\alpha_i - \hat{\alpha}_{i|j})(\alpha_i - \hat{\alpha}_{i|j})'\}$$

حال فرض کنید که مقدار سری زمانی در لحظه t نیز مشاهده شود؛ با استفاده از روابط زیر برآورد حالت و ماتریس کوواریانس بردار خطای حالت به روز رسانی می‌شود:

$$\hat{\alpha}_{t|t} = \hat{\alpha}_{t|t-1} + K_t (X_t - H_t \hat{\alpha}_{t|t-1}) \quad (8)$$

$$P_{t|t} = [I - K_t H_t] P_{t|t-1} \quad (9)$$

که در آن ماتریس K_t را ماتریس حاصل کالمن می‌نامند و به صورت زیر تعریف می‌شود:

$$K_t = P_{t|t-1} H'_t [H_t P_{t|t-1} H'_t + Q_t]^{-1} \quad (10)$$

به این ترتیب با در نظر گرفتن شرایط آغازین زیر، می‌توان پیش‌بینی را به صورت بازگشتی انجام داد.

$$\hat{\alpha}_0 = E[\alpha_0]$$

$$P_0 = E\{(\alpha_0 - E(\alpha_0))(\alpha_0 - E(\alpha_0))'\}$$

۲-۲ فیلتر کالمن تعمیم یافته

فیلتر کالمن که در بخش پیش معرفی شد، به دلیل استفاده از عملگرهای خطی تنها برای مدل‌های فضای حالت خطی (روابط ۳ و ۴) قابل استفاده است. برای مدل‌های غیر خطی (روابط ۱ و ۲) با اضافه کردن یک فرآیند خطی سازی می‌توان فیلتر کالمن را تعمیم داد (هایکین ۲۰۰۱). به این ترتیب می‌توان فیلتر کالمن تعمیم یافته را فرمول‌بندی کرد. فرض کنید که سری زمانی X_t که با معادلات (۱) و (۲) نمایش داده شد، در زمان $t-1$ قرار داشته باشد و بخواهیم وضعیت سری زمانی و حالت فرآیند مارکف را برای لحظه t پیش‌بینی کنیم. دو ماتریس زیر را تعریف می‌کنیم:

$$F_{t-1,t} = \frac{\partial f(t-1, \alpha)}{\partial \alpha} \Big|_{\alpha=\hat{\alpha}_{t-1|t-1}} \quad (11)$$

$$H_t = \frac{\partial h(t, \alpha)}{\partial \alpha} \Big|_{\alpha=\hat{\alpha}_{t|t-1}} \quad (12)$$

حال با استفاده از روابط (۱۱) و (۱۲) معادلات پیش‌بینی عبارت خواهند بود از:

$$\hat{\alpha}_{t|t-1} = f(t, \hat{\alpha}_{t-1|t-1}) \quad (13)$$

$$P_{t|t-1} = F_{t-1,t} P_{t-1|t-1} F'_{t-1,t} + R_{t-1} \quad (14)$$

$$\hat{X}_t = h(t, \hat{\alpha}_{t|t-1}) \quad (15)$$

با اضافه شدن مشاهده t ام، معادلات به روز رسانی را نیز می‌توان به صورت زیر بازنویسی کرد:

$$\hat{\alpha}_{t|t} = \hat{\alpha}_{t|t-1} + K_t X_t - h(t, \hat{\alpha}_{t|t-1}) \quad (16)$$

$$P_{t|t} = [I - K_t H_t] P_{t|t-1} \quad (17)$$

و ماتریس حاصل کالمن نیز عبارت است از:

$$K_t = P_{t|t-1} H'_t [H_t P_{t|t-1} H'_t + Q_t]^{-1} \quad (18)$$

به این ترتیب با در اختیار داشتن ماتریس ضرایب و فرم‌های تابعی f و h می‌توان به پیش‌بینی سری زمانی و حالت فرآیند مارکف پنهان پرداخت.

۳ نمایش فضای حالت و معادلات پیش‌بینی برای شبکه‌های عصبی

یک مدل شبکه عصبی به صورت زیر را در نظر بگیرید:

$$y_t = g(x_t) + w_t \quad (19)$$

که در آن $w_t \sim N(0, \sigma^2)$ و $x_t = [y_{t-1}, y_{t-2}, \dots, y_{t-p}]'$ است. همچنین یک بردار حالت به صورت $\alpha_t = [y_{t-p+1}, \dots, y_{t-1}, y_t]$ را در نظر بگیرید. مدل (۱۹) را می‌توان به صورت یک مدل فضای حالت نشان داد:

$$\alpha_{t+1} = F(\alpha_t) + w_{t+1} \quad (20)$$

$$y_t = H\alpha_t \quad (21)$$

که در آن:

$$F(\alpha_t) = [y_{t-p+2}, \dots, y_t, g(\alpha_t)]'$$

$$w_t = w_t[0, \dots, 0, 1]'$$

$$H = [0, \dots, 0, 1]$$

اگر $w_t \sim N(0, \Sigma_w)$ با استفاده از فیلتر کالمن تعمیم یافته می‌توان معادلات پیش‌بینی را به صورت زیر نوشت:

$$\hat{y}_{t+h} = H\hat{\alpha}_{t+h} \quad (22)$$

$$\hat{\alpha}_{t+j} = \hat{F}(\hat{\alpha}_{t+j-1}) \quad (23)$$

$$\hat{P}_{t+j|t+j-1} = \hat{f}_{t+j-1} \hat{P}_{t+j-1|t+j-1} \hat{f}'_{t+j-1} + \Sigma_w \quad (24)$$

که در آن:

$$f_t = \left. \frac{\partial F(\alpha)}{\partial \alpha} \right|_{\alpha=\alpha_t}$$

با اضافه شدن مشاهدات جدید، برای به روز رسانی از معادلات زیر استفاده می‌شود.

$$\alpha_{t+j} = \hat{\alpha}_{t+j} + K_{t+j}(y_{t+j} - H\hat{\alpha}_{t+j}) \quad (25)$$

$$\hat{P}_{t+j|t+j} = [I - K_{t+j}H] \hat{P}_{t+j|t+j-1} \quad (26)$$

$$K_{t+j} = \hat{P}_{t+j|t+j-1} H' [H\hat{P}_{t+j|t+j-1} H']^{-1} \quad (27)$$

۱-۳ معادلات پیش‌بینی برای شبکه عصبی چند لایه خود بازگشت

فرض کنید برای مدل سازی سری زمانی y_t از یک شبکه عصبی پیش‌خور با یک لایه پنهان به صورت زیر استفاده کنیم:

$$y_t = g(\mathbf{x}_t) + w_t = \gamma_0 + \sum_{i=1}^m \gamma_i h \left(\beta_{i0} + \sum_{j=1}^p \beta_{ij} y_{t-j+1} \right) + w_t \quad (28)$$

که در آن m تعداد نرون‌های لایه پنهان، $\mathbf{x}_t = [y_{t-1}, y_{t-2}, \dots, y_{t-p}]'$ و h یک تابع فعالیت مشتق‌پذیر است. در این صورت با استفاده از رابطه (۲۰) و (۲۱)، مدل (۲۸) را می‌توان به صورت یک مدل فضای حالت نوشت. در این صورت با استفاده از روابط (۲۲)، (۲۳) و (۲۴) معادلات پیش‌بینی برای شبکه عصبی پیش‌خور با یک لایه پنهان به صورت زیر خواهد بود:

$$\hat{y}_{t+h} = H \hat{\alpha}_{t+h} \quad (29)$$

$$\hat{\alpha}_{t+j} = \hat{F}(\hat{\alpha}_{t+j-1}) \quad (30)$$

$$\hat{P}_{t+j|t+j-1} = \hat{f}_{t+j-1} \hat{P}_{t+j-1|t+j-1} \hat{f}'_{t+j-1} + \Sigma_w \quad (31)$$

اگر $h(x)$ یک تابع لجستیک به صورت زیر باشد

$$h(x) = \frac{1}{1 + e^{-cx}}$$

در این صورت:

$$\mathbf{f}_t = \frac{\partial \mathbf{F}(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \Big|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_t} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ \frac{\partial g(\mathbf{x}_t)}{\partial y_{t-p+1}} & \frac{\partial g(\mathbf{x}_t)}{\partial y_{t-p+2}} & \frac{\partial g(\mathbf{x}_t)}{\partial y_{t-p+3}} & \dots & \frac{\partial g(\mathbf{x}_t)}{\partial y_t} \end{pmatrix},$$

$$\frac{\partial g(\mathbf{x}_t)}{\partial y_{t-k+1}} = \sum_{i=1}^m \left[\gamma_i \beta_{ik} h \left(\beta_{i0} + \sum_{j=1}^p \beta_{ij} y_{t-j+1} \right) \left(1 - h \left(\beta_{i0} + \sum_{j=1}^p \beta_{ij} y_{t-j+1} \right) \right) \right]$$

با اضافه شدن مشاهدات جدید، برای به روز رسانی نیز از معادلات (۲۵)، (۲۶) و (۲۷) می‌توان استفاده کرد.

۲-۳ مقادیر آغازین

برای استفاده از معادلات بازگشتی برآورد بردار حالت، باید برآوردی از مقادیر اولیه بردار حالت و واریانس اولیه آن بدست آورد. اگر مدل فضای حالت، یک مدل ایستا و زمان همگن باشد در این صورت می‌توان به سادگی از برآورد میانگین و واریانس غیر شرطی بردار حالت به عنوان برآورد مقادیر اولیه بردار حالت و واریانس اولیه آن استفاده کرد (آکائیک ۱۹۷۸). با توجه به اینکه مدل‌های شبکه عصبی تحت شرایطی مجانباً ایستا هستند و در این شرایط میانگین و واریانس بردار حالت در طول زمان تغییر می‌کنند، نمی‌توان از برآورد میانگین و واریانس غیر شرطی برای این مدل‌ها استفاده کرد. روش‌های مختلفی برای برآورد مقادیر اولیه بردار حالت و واریانس اولیه آن پیشنهاد شده است. به عنوان نمونه هاروی و فیلیپس (۱۹۷۹) با این استدلال که واریانس بزرگ باعث می‌شود تأکید روی مقادیر خاص مقدار اولیه بردار حالت وجود نداشته باشد، یک واریانس بسیار بزرگ را برای مقدار اولیه بردار حالت پیشنهاد کرده‌اند. به عبارت دیگر فرض اساسی در این روش آن است که مقدار اولیه بردار حالت دارای یک توزیع ناآگاهی بخش^۲ است. کیتاگاوا (۱۹۸۱) نیز با استفاده از این فرض که معکوس واریانس مقدار اولیه بردار حالت به صفر میل می‌کند، روشی را برای برآورد مقادیر اولیه ارائه کرد. همچنین دی جانگ (۱۹۸۸) پیشنهاد کرد از یک روش بازگشتی برای برآورد مقادیر اولیه استفاده شود. روش پیشنهادی هاروی و فیلیپس (۱۹۷۹) بیشتر مورد استناد قرار گرفته زیرا حساسیت فیلتر کالمن به مقایسه اولیه را تا حد زیادی از بین می‌برد و در عمل نیز به سادگی قابل پیاده سازی است. به این ترتیب فرض می‌کنیم واریانس اولیه بردار حالت به صورت زیر فرمول‌بندی می‌شود:

$$P_0 = kI$$

که در آن I ماتریس همانی با بعد $p \times p$ است و k یک مقدار مثبت که $k \rightarrow \infty$ (در عمل، به دلیل محدودیت در حافظه سیستم‌های محاسباتی، k باید یک مقدار تا حد امکان بزرگ انتخاب شود). برای برآورد بردار حالت، با توجه به اینکه تمام عناصر بردار حالت در مدل (۲۰) و (۲۱) قابل مشاهده هستند، پس از چند مرحله تکرار در فیلتر کالمن، اثر مقادیر اولیه

²Noninformative

کاملاً حذف خواهد شد. به بیان دیگر معادلات پیش‌بینی و به روز رسانی ارائه شده در بخش قبل، نسبت به مقدار اولیه فیلتر کالمن حساس نیستند. بنا بر این به عنوان یک برآورد می‌توان از میانگین p مشاهده اول استفاده کرد. به این ترتیب میانگین و واریانس مقدار اولیه بردار حالت را می‌توان به صورت زیر برآورد کرد:

$$\hat{E}(\alpha_0) = \frac{1}{p} \sum_{i=1}^p y_i$$

$$\hat{Var}(\alpha_0) = P_0 = kI$$

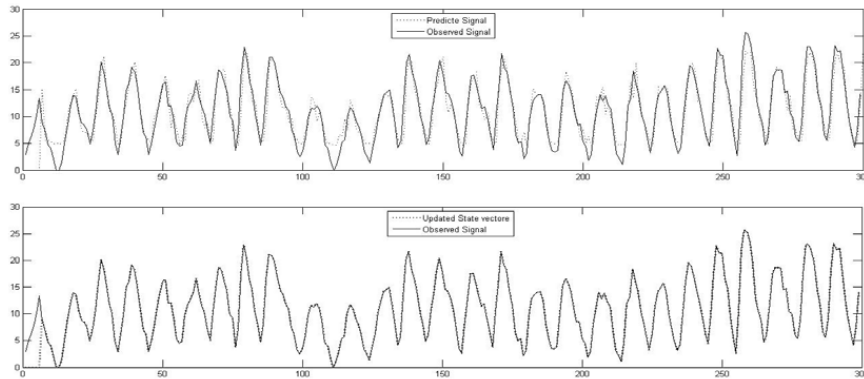
۴ مثال عددی

برای بررسی توانایی روش ارائه شده در بخش قبل، معادلات پیش‌بینی و به روز رسانی برای شبکه‌های عصبی خود بازگشت، در این بخش از روش فوق برای پیش‌بینی سری لکه‌های خورشیدی استفاده شده است. داده‌های مورد بررسی، مشاهدات مربوط به لکه‌های خورشیدی سالانه از سال ۱۷۰۰ تا ۱۹۹۸ است که در پایگاه اینترنتی مرکز ملی داده‌های ژئوفیزیک^۳ منتشر شده است. برای برازش مدل شبکه‌عصبی خودبازگشت و انجام پیش‌بینی از تبدیل پیشنهاد شده توسط قدار و تانگ (۱۹۸۱)، $y_t = 2[\sqrt{1 + N_t} - 1]$ ، استفاده شده است که در آن N_t تعداد لکه‌های خورشیدی مشاهده شده در طول یک سال است. مدل شبکه عصبی برازش شده دارای شش گام وابستگی به عنوان ورودی‌های مدل و دو نرون در لایه پنهان است:

$$y_t = \gamma_0 + \sum_{i=1}^2 \gamma_i h \left(\beta_{i0} + \sum_{j=1}^6 \beta_{ij} y_{t-j+1} \right) + w_t.$$

³www.ngdc.noaa.gov

نمودار (۱)، مقادیر مشاهده شده، مقادیر پیش‌بینی و مقادیر به روز رسانی بردار حالت، α_t ، را نشان می‌دهد.



شکل ۱: مقادیر مشاهده شده، مقادیر پیش‌بینی و مقادیر به روز رسانی بردار حالت با مقایسه مقادیر به روز شده بردار حالت و مقادیر مشاهده شده سری در نمودار پایین، مشاهده می‌شود که مقادیر به روز رسانی تنها در چند تکرار اول از مقدار آغازین اثر می‌پذیرند و به سرعت با سری اصلی انطباق پیدا می‌کنند. ماتریس کوواریانس به روز رسانی شده بردار حالت نیز این واقعیت را تأیید می‌کند. مقادیر ماتریس کوواریانس به روز رسانی شده از تکرار پانزدهم به بعد تقریباً صفر است. برآورد واریانس مقدار برآورد شده بردار حالت نیز برابر ۲/۳۵ است که با واریانس برآورد شده برای خطای پیش‌بینی مطابقت دارد. جدول ۱ مقادیر مشاهده شده، مقادیر پیش‌بینی و به روز رسانی سری لکه‌های خورشیدی بین سال‌های ۱۹۸۹ تا ۱۹۹۸ را نشان می‌دهد.

۵ نتیجه‌گیری

در این مقاله روشی جدید برای پیش‌بینی در مدل‌های شبکه عصبی خودبازگشت، با استفاده از فیلتر کالمن تعمیم یافته، ارائه شد. برای بررسی توانایی این روش، معادلات به دست آمده برای پیش‌بینی سری لکه‌های خورشیدی مورد استفاده قرار گرفتند. نتایج نشان می‌دهد در این روش، به دلیل رابطه بردار حالت و سری قابل مشاهده، در شرایطی که اندازه‌گیری‌ها با خطا همراه نباشند، بردار حالت به روز رسانی شده و پیش‌بینی آن حساسیت بسیار کمی

جدول ۱: مقادیر مشاهده شده، پیش‌بینی و به روز رسانی سری لکه‌های خورشیدی برای سال‌های ۱۹۸۹ تا ۱۹۹۸.

بردار حالت به هنگام شده	پیش‌بینی	مشاهده	
۱۰۰/۲	۱۲۷/۹	۱۰۰/۲	۱۹۸۹
۱۵۷/۶	۱۳۹/۴	۱۵۷/۶	۱۹۹۰
۱۴۵/۷	۱۲۰/۲	۱۴۵/۷	۱۹۹۱
۹۴/۳	۱۱۴/۷	۹۴/۳	۱۹۹۲
۵۴/۶	۵۶/۱	۵۴/۶	۱۹۹۳
۲۹/۹	۳۰/۸	۲۹/۹	۱۹۹۴
۱۷/۵	۱۹/۲	۱۷/۵	۱۹۹۵
۸/۶	۱۲/۶	۸/۶	۱۹۹۶
۲۱/۵	۱۲/۸	۲۱/۵	۱۹۹۷
۶۴/۳	۴۹/۶	۶۴/۳	۱۹۹۸
	۱۵۳/۵		۱۹۹۹

نسبت به مقادیر اولیه نشان می‌دهند. همچنین به شرط کارایی مدل شبکه‌عصبی، معادلات پیش‌بینی نتایج مناسبی ارائه می‌کنند. استفاده از فضای حالت و فیلتر کالمن تعمیم یافته، امکان انجام پیش‌بینی حتی در حالتی که سری مورد بررسی ایستا و یا گوسی نباشد را ایجاد می‌کند. علاوه بر این، می‌توان با استفاده از معادلات ارائه شده، در شرایطی که اندازه‌گیری با خطا انجام می‌گیرد، مدل مورد بررسی را تعمیم داد (هایکین ۲۰۰۱).

مراجع

- [1] Akaike H. (1978). Time series analysis and control through parametric model, in Findley (Eds.) *Applied Time Series Analysis*, Academic Press.
- [2] De Jong P. (1988). A Cross-validation filter for time series models, *Biometrika*, **75**, 594-600.
- [3] Franses P.H. and van Dijk D. (2000). *Non-linear Time Serise Models in Empirical Finance*, Cambridge University Press, Cambridge.
- [4] Ghaddar D.K. and Tong H. (1981). Data Transformation and self-exciting threshold Autoregression, *Journal of Royal Statistical Society*, **C30**, 238-248.
- [5] Harvey A.C. and Phillips G.D.A. (1979). Maximum likelihood estimation of regression models with autoregressive-moving average disturbances, *Biometrika* **66**, 49-58.
- [6] Haykin S. (2001). Kalman filters, in Haykin S. (Ed.), *Kalman Filtering and Neural Networks*, John Wiley & Sons, New York, 1-21.
- [7] Kalman R.E. (1960). A new approach to linear filtering and prediction problems, *Transactions of ASME, Ser. D, Journal of Basic Engineering*, **82**, 34-45.
- [8] Kalman R.E. and Bucy R.S. (1961). New results in filtering and prediction theory, *Trans. ASME Journal of Basic Engineering*, **83**, 95-108.

- [9] Kitagawa G. (1981). A nonstationary time series model and its fitting by recursive technique, *Journal of Time Series Analysis*, **2**, 103-116.
- [10] Plumer E.S. (1995). Training neural networks using sequential extended Kalman filtering, *Proceedings of World Congress on Neural Networks*, I-764 - I-769.
- [11] Priestly M.B. (1988). *Nonlinear and Non-stationary Time Series Analysis*, Academic Press, London.
- [12] Puskorius G.V. and Feldkamp L.A. (2001). Parameter based Kalman filter training: theory and implementation, in Haykin S. (Ed.), *Kalman Filtering and Neural Networks*, John Wiley & Sons, New York, 23-67.
- [13] Puskorius G.V. and Feldkamp L.A. (1997). Extensions and enhancements of decoupled extended Kalman filter training, *Proceedings of International Conference of Neural Networks*, **3**, 1879-1883.
- [14] Puskorius G.V. and Feldkamp L.A. (1991). Decoupled extended Kalman filter training of feedforward layered networks, *Proceedings of International Joint Conference of Neural Networks*, **1**, 771-777.
- [15] Singal S. Wu L. (1989). Training multilayer perceptrons with the extended Kalman algorithm, in Touretzky D.S., Eds., *Advances in Neural Information Processing systems 1*, Morgan Kaufmann, 133-140.
- [16] Sun P. and Marko K. (1998). The square root Kalman filter training of recurrent neural networks, *Proceedings of the IEEE Conference on Systems, Man and Cybernetics*, 1645-1651.