بسم الله الرحمن الرحیم

دانشگاه صنعتی امیر کبیر

(پلی تکنیک تهران)

دانشکده ریاضی و علوم کامپیوتر

پایان نامه کارشناسی ارشد

رشته : آمار ریاضی

عنوان :

مدلبندی توزیع انتقال آمیخته سریهای زمانی با واریانس ناپایدار

نگارنده : شبنم فانی

استاد راهنما : دکتر سعید رضاخواه

استاد مشاور : دکتر عادل محمدپور

مهر ۱۳۸٦

دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)
معاونت پژوهشی

بسمه تعالی

فرم اطلاعات پایان نامه
کارشناسی ارشد و دکترا

تاریخ: ....
پیوست: ....

نام و نام خانوادگی:	شبنم فانی	**دانشجوی آزاد** ■		**بورسیه** ☐		**معادل** ☐	
شماره دانشجوئی:	۸۴۱۱۳۰۰۹	**دانشکده:**	ریاضی و علوم کامپیوتر	**رشته تحصیلی:**	آمار		

نام و نام خانوادگی استاد راهنما: سعید رضاخواه

عنوان پایان نامه به فارسی: مدلبندی توزیع انتقال آمیخته سریهای زمانی با واریانس ناپایدار

عنوان پایان نامه به انگلیسی: Mixture transition distribution (MTD) modeling of herteroscedastic time series

نوع پروژه: کاربردی ☐ بنیادی ☐ توسعه ای ☐ نظری ■

تاریخ شروع: اسفند ۸۵ **تاریخ خاتمه:** مهر ۸۶ **تعداد واحد:** ۶

سازمان تأمین کننده اعتبار: مرکز تحقیقات و مخابرات ایران

واژه های کلیدی به فارسی: سریهای زمانی، ناپایداری واریانس یا ناهم واریانسی، آمیخته، مدل HMTD، الگوریتم EMتعمیم یافته، تباهیدگی لگاریتم درستنمایی

واژه های کلیدی به انگلیسی: Time series, Heteroscedasticity, Mixture, HMTD, GEM algorithm,Dwgenerescence of the log-likelihood

نظرها و پیشنهادها به منظور بهبود فعالیت های پژوهشی دانشگاه:

استاد راهنما:

دانشجو:

امضاء استاد راهنما: **تاریخ:**

نسخه ۱: معاونت پژوهشی

نسخه ۲: کتابخانه و به انضمام دو جلد پایان نامه به منظور تسویه حساب با کتابخانه و مرکز اسناد و مدارک علمی

به پاس عاطفه سرشار و گرمای امیدبخش وجودشان که در این سردترین روزگاران بهترین پشتیبان است

به پاس قلب‌های بزرگشان که فریاد رس است و سرگردانی و ترس در پناهشان به شجاعت می‌گراید

به پاس تعبیر عظیم و انسانی‌شان از کلمه ایثار و از خود گذشتگان

و به پاس محبت‌های بی‌دریغشان که هرگز فروکش نمی‌کند

این پایان نامه را به پدر و مادر عزیزم تقدیم می‌کنم.

تقدیر و تشکر

اینک که به حول و لطف الهی تحصیلات دوره کارشناسی ارشد را به پایان می‌رسانم، سزاوار است تشکر قلبی خود را نثار عزیزانی کنم که من را در انجام این مشکل، یاری داده‌اند. اساتید بزرگواری که با بذل وجود خود، اندوخته ادبی و علمی اینجانب را تا این حد ارتقا دادند.

با تشکر فراوان از استاد گرانقدر جناب آقای دکتر سعید رضاخواه که من را در انجام این پایان نامه یاری رساندند و از هیچ کوششی در این مسیر دریغ ننموده‌اند. همچنین از جناب آقای دکتر عادل محمدپور جهت راهنمایی‌های ارزشمندشان کمال تشکر را دارم.

سرانجام از تمام دوستان عزیزی که از جهات مختلف مرا همراهی و مساعدت کرده و با راهنمایی‌های خود زمینه را برای فعالیت بهتر اینجانب مهیا نمودند، قدردانی می‌کنم.

چکیده

بسیاری از سری‌های زمانی دارای ویژگی‌های غیرنرمال هستند ازجمله سری‌های زمانی با واریانس ناپایدار[1]. این نوع از سری‌های زمانی نقش محوری در بسیاری از علوم مانند ریاضیات مالی، بیمه و زلزله‌شناسی ایفا می‌کنند. مدل توزیع انتقال آمیخته [2]، اولین بار توسط رفتری [3] در سال 1985 برای مدلبندی زنجیرهای مارکف مرتبه بالا معرفی شد.

در این پایان نامه، مدل توزیع انتقال آمیخته در مورد سری‌های زمانی با واریانس ناپایدار *HMTD* [4] به عنوان تعمیمی از این نوع سری‌های زمانی مورد مطالعه قرار گرفته شده است. در این نوع مدل‌ها با در نظر گرفتن امید ریاضی و انحراف معیار هر مؤلفه برحسب تابعی از مشاهدات گذشته فرآیند، مدلی مناسب از توزیع احتمال شرطی مشاهدات آینده بدست می‌آید. با استفاده از یک مثال عددی نشان داده خواهد شد که مدل *HMTD* می‌تواند بهتر از مدل‌هایی مانند *ARMA* و *GARCH* [5] عمل کند. همچنین نتایج متفاوتی از برآورد پارامترهای مدل آمیخته مطرح خواهد شد. مقاله اصلی استفاده شده در این پایان نامه بر پایه [4] است که در قسمت مراجع ذکر شده است.

کلمات کلیدی: سری های زمانی[6]، ناپایداری واریانس یا ناهم‌واریانسی، آمیخته [7]، مدل *HMTD*، الگوریتم *EM* تعمیم‌یافته[8] *GEM*، تباهیدگی لگاریتم درستنمایی[9]

---

[1] Heteroscedasticity
[2] Mixture Transition Distribution Model
[3] Raftery
[4] Heteroscedastic Mixture Transition Distribution Model
[5] General Auto Regressive Conditional Heteroscedasticity
[6] Time series
[7] Mixture
[8] Generalized Expectation-Maximization Algorithm
[9] Degenerescence of the Log-Likelihood

فهرست مطالب

فصل اول: مفاهیم پایه‌ای

فصل دوم: معرفی مدل توزیع انتقال آمیخته

فصل سوم: معرفی الگوریتم *EM* و ژنتیک

فهرست اشکال

فهرست جدول

مقدمه

در بررسی داده‌های سری زمانی معمولاً فرض ثابت بودن واریانس را بکار می‌گیرند. اما در عمل گاهی واریانس ثابت نیست. از این رو می‌بایست داده‌ها را به‌نحوی تبدیل نماییم تا پایداری در واریانس حاصل شود یا از مدل‌هایی استفاده کنیم که مشروط به ناپایداری در واریانس هستند. یک خانواده معروف از این مدل‌ها، خانواده مدل‌های مشروط به ناپایداری واریانس می‌باشد که به نام مدل‌های اتورگرسیو شرطی با واریانس ناپایدار[1] ($ARCH$) معروف است و برای اولین بار در سال ۱۹۸۲ توسط انگل[2] در [12] معرفی شد. پس از آن در سال ۱۹۸۶ شخص دیگری به نام بولرسلو[3]، مدل دیگری به نام مدل تعمیم‌یافته اتورگرسیو شرطی با واریانس ناپایدار[4] ($GARCH$) در [7] را بیان کرد که در آن واریانس علاوه بر توان دوم مشاهدات، برحسب واریانس‌های گذشته نیز بیان می‌شود.

از طرفی رفتری در سال ۱۹۸۵ مدل توزیع انتقال آمیخته را به منظور مدلبندی زنجیرهای مارکف مرتبه بالا با فضای وضعیت متناهی معرفی کرد. سپس در سال ۱۹۹۶ لی، مارتین و رفتری[5] دو تعمیم از این مدل‌ها با فضای وضعیت دلخواه را ارائه کردند. در این دو تعمیم تنها امیدریاضی هر مؤلفه مدل آمیخته برحسب تابعی از مشاهدات گذشته فرآیند بیان می‌شود و واریانس هر مؤلفه ثابت خواهد بود. پس از آن برچتلد[6] در سال ۲۰۰۳ تعمیم دیگری از این خانواده مدل‌ها به نام مدل توزیع انتقال آمیخته در مورد سری‌های زمانی با واریانس ناپایدار را ارائه کرد که در آن علاوه بر اینکه امیدریاضی هر مؤلفه مدل آمیخته برحسب گذشته فرآیند بیان می‌شود، انحراف معیار نیز به عنوان تابعی از گذشته فرآیند در نظر گرفته می‌شود و این مدل را مدل توزیع انتقال آمیخته سری‌های زمانی با واریانس ناپایدار ($HMTD$) نامید که این مدل توانایی بررسی رفتارهای متنوعی از سری‌های غیرنرمال را دارد. هدف ما از مطالعه این مدل بدست آوردن برآورد تابع توزیع احتمال شرطی مشاهده آینده به‌جای استفاده از برآورد نقطه‌ای می‌باشد.

بنابراین با توجه به مطالب بیان شده، در فصل اول مفاهیم پایه‌ای مورد نیاز در بخش‌های بعدی را بیان می‌کنیم. در فصل دوم مدل توزیع انتقال آمیخته را معرفی کرده و دو تعمیم از این مدل (مدل توزیع

---
[1] Auto Regressive Conditional Heteroscedasticity
[2] Engle
[3] Bollerslev
[4] General Auto Regressive Conditional Heteroscedasticity
[5] Le, Martin & Raftery
[6] Berchtold

انتقال آمیخته نرمال $GMTD$ و مدل اتورگرسیو آمیخته MTD) را مورد مطالعه قرار می‌دهیم. در فصل سوم به منظور برآورد پارامترهای این دو مدل تعمیم یافته، الگوریتم ژنتیک و EM را معرفی می‌کنیم. در فصل چهارم مدل توزیع انتقال آمیخته سری‌های زمانی با واریانس ناپایدار ($HMTD$) را معرفی و مورد مطالعه قرار می‌دهیم. در عمل با توجه به مثال بیان شده در این زمینه، نشان داده خواهد شد که این مدل برازش بهتری از سری‌های زمانی را نسبت به مدل‌های $GARCH$، $GMTD$ و MAR ارائه می‌دهد. برآورد پارامترهای مدل $HMTD$ با استفاده از الگوریتم ژنتیک و EM تعمیم‌یافته بدست می‌آید که بدین منظور از بسته نرم‌افزاری $MATLB$ 7,3 استفاده شده است. مقاله اصلی استفاده شده در این پایان‌نامه توسط برچتلد نوشته شده که در قسمت مراجع [4] ذکر شده است.

فصل اول

تعاریف و مفاهیم پایه‌ای

در این فصل مفاهیم کلی و مقدمات مورد نیاز فصول بعد ذکر شده است.

تعریف ۱-۱ فرآیند تصادفی

یک فرآیند تصادفی، خانواده‌ای از متغیرهای تصادفی $X(w,t)$ است، که در آن w به فضای نمونه و t به یک مجموعه اندیس‌گذار، متعلق می‌باشد. برای یک ثابت t، $X(w,t)$ یک متغیر تصادفی است. برای یک w معلوم، $X(w,t)$ به عنوان تابعی از t، یک تابع نمونه یا یک مصداق نامیده می‌شود. جامعه‌ای متشکل از تمام مصادیق ممکن در فرآیندهای تصادفی و تحلیل سری‌های زمانی، یک مجموعه کلی نامیده می‌شود. بنابراین یک سری زمانی، یک مصداق یا یک تابع نمونه از یک فرآیند تصادفی معین است. $X(w,t)$ به عنوان مجموعه‌ای از متغیرهای تصادفی در یک فضای نمونه تعریف می‌شود و ما معمولاً $X(w,t)$ را به‌صورت $X(t)$ یا X_t می‌نویسیم.

هر فرآیند تصادفی با دو عامل اصلی فضای وضعیت S و مجموعه اندیسگذار T شناخته می‌شود.

فضای وضعیت S:

فضایی است که مقادیر ممکن هر X_t در آن قرار می‌گیرند. در حالتی‌که $S=\{0,1,...\}$ باشد، فرآیند مربوطه را با مقدار صحیح یا یک فرآیند با وضعیت گسسته می‌نامیم. هرگاه S خط حقیقی $(-\infty,\infty)$ باشد، آنگاه X_t را یک فرآیند تصادفی با مقدار حقیقی می‌نامیم. هرگاه S یک فضای اقلیدسی k-بعدی باشد، آنگاه گوئیم X_t یک فرآیند برداری k-بعدی است.

مجموعه اندیسگذار T:

هرگاه T زیرمجموعه‌ای از اعداد صحیح باشد، آن‌گاه X_t یک فرآیند تصادفی با زمان گسسته است و هرگاه T فاصله‌ای با طول مثبت باشد، فرآیند را فرآیندی با زمان پیوسته می‌نامند.

تعریف ۱-۲ سری زمانی

مجموعه‌ای از مشاهدات متوالی است که بر حسب زمان مرتب یا ضبط شده‌اند.

تعریف ۱-۳ فرآیند $i.i.d$

ساده‌ترین سری زمانی است که بدون روند[1]، بدون نمود فصلی و بدون ناپایداری در واریانس است. این سری مشاهدات مستقل و هم‌توزیع هستند.

تعریف ۱-٤ تابع محدب

تابع پیوسته $g(.)$ که قلمرو و مجموعهٔ تصویر آن اعداد حقیقی است، محدب نامیده می‌شود اگر برای هر x_0 بر مجموعه اعداد حقیقی، خطی وجود داشته باشد که از نقطهٔ $(x_0,g(x_0))$ بگذرد و بر نمودار تابع یا زیر نمودار تابع $g(.)$ قرار گیرد.

[1] trend

تعریف ۱-۵ نامساوی جنسن

فرض کنید X یک متغیر تصادفی با میانگین $E[X]$ باشد. اگر $g(.)$ تابعی محدب با $E[g(X)]$ متناهی باشد، آنگاه

$$E[g(X)] \geq g(E[X])$$

در صورتی‌که $g(.)$ تابعی مقعر باشد، جهت نامساوی بالا تغییرمی‌کند (اثبات در [29]).

تعریف ۱-۶ فرمول بیز

برای فضای احتمال $(\Omega, F, P[.])$، اگر $B_1, B_2, \ldots, B_n$ مجموعه‌ای از پیشامدهای دوبه‌دو جدا در F باشند که در شرایط زیر صدق کنند:

برای $\Omega = \bigcup_{j=1}^{n} B_j$ و $P[B_j] > 0,\ j = 1, \ldots, n$

آنگاه برای هر $A \in F$ که در $P[A] > 0$ صدق کند، داریم:

$$P[B_k | A] = \frac{P[A|B_k]\, P[B_k]}{\sum_{j=1}^{n} P[A|B_j]\, P[B_j]}$$

(اثبات آن در [29] اشاره شده است).

تعریف ۱-۷ فرآیند $ARMA(p,q)$

X_t یک فرآیند $ARMA(p,q)$ است اگر X_t ایستا بوده و برای هر t داشته باشیم:

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$$

به‌طوریکه $\{Z_t\} \sim WN(0, \sigma^2)$ و چندجمله‌ای‌های $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$ ، $\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$ هیچ فاکتور مشترکی ندارند. نمایش $ARMA(p,q)$ بر اساس این چندجمله‌ای‌ها به‌صورت زیر می‌باشد:

$$\phi(B)\,X_t=\theta(B)\,Z_t$$

و B عملگر برگشت به عقب می‌باشد $\left(B^j X_t = X_{t-j}; B^j Z_t = Z_{t-j}\right)$.

اگر $\theta(z)=1$ باشد، سری زمانی X_t را فرآیند اتورگرسیو مرتبه p یا $AR(p)$ می‌گویند و اگر $\phi(z)=1$ باشد، سری زمانی X_t را فرآیند میانگین متحرک مرتبه q یا $MA(q)$ می‌گویند.

تعریف ۱–۸ فرآیندهای ایستا

یک فرآیند ایستا فرآیندی تصادفی است که قوانین احتمالی‌اش در طی انتقالاتی در زمان بدون تغییر می‌باشند. این مفهوم ناشی از مفهوم بسیار طبیعی یک دستگاه فیزیکی است که فاقد مبدأ ذاتی زمان می‌باشد و برای بسیاری از فرآیندها در نظریه ارتباطات، ستاره‌شناسی، زیست‌شناسی و اقتصاد فرض مناسبی خواهد بود.

تعریف ۱–۹ σ– میدان

کلاس F از زیر مجموعه‌های Ω یک σ– میدان نامیده می‌شود هرگاه

(i) $\Omega \in F$

(ii) اگر $A \in F$ باشد نتیجه شود که $A^c \in F$

(iii) اگر $A_1, A_2, \ldots \in F$ باشد ، آنگاه نتیجه شود که $A_1 \cup A_2 \cup \ldots \in F$.

تعریف ۱–۱۰ تابع درستنمایی

فرض کنید $X=(x_1,\ldots,x_n)$ بردار n متغیر تصادفی با تابع چگالی احتمال توأم (یا تابع احتمال توأم) $f_\theta(x)$ ، $\theta \in \Theta \subseteq R^k$، باشد. برای هر مقدار داده شده $X=x$، تابع درستنمایی X را تابع چگالی احتمال توأم X، یعنی $f_\theta(x)$ ، تعریف می‌کنیم که به‌صورت تابعی از θ در نظر گرفته می‌شود و آن را با نماد $L(\theta)$ نمایش می‌دهیم:

$$L(\theta)=f_\theta(x)=L(\theta;x)$$

تعریف ۱–۱۱ برآوردگر ماکسیمم درستنمایی

اگر $\delta(X)$ برآوردگری برای θ باشد، به‌طوریکه

$$P_\theta(\delta(X)\in\Theta)=1;\forall\theta\in\Theta \text{ (i)}$$

$$L(\delta(X))\geq L(\theta);\forall\theta\in\Theta \text{ (ii)}$$

آنگاه $\delta(X)$ به عنوان برآوردگر درستنمایی ماکسیمم θ تعریف می‌شود.

تعریف ۱–۱۲ مدل اتورگرسیو شرطی با واریانس ناپایدار

فرآیند $\{r_t\}$ که بر اساس معادلات

$$r_t=\sigma_t\varepsilon_t \qquad t=1,2,\ldots$$
$$\sigma_t^2=a+b_1r_{t-1}^2+\cdots+b_pr_{t-p}^2$$

معرفی می‌شود و $\{\varepsilon_t\}\sim IID(0,1)$ بوده و ε_t از $\{r_{t-k},k\geq1\}$ مستقل می‌باشد را یک فرآیند اتورگرسیو شرطی با واریانس ناپایدار[2] $ARCH(p)$ می‌نامند. مدل اتورگرسیو شرطی با پراکندگی ناپایدار ($ARCH$) که توسط انگل [12] در سال ۱۹۸۲ معرفی گردید، یک تابع خطی از توان دوم گذشته سری است.

در این مدل برای اینکه از مثبت بودن واریانس مطمئن شویم باید قیدهای زیر در مورد پارامترها برقرار باشد:

$$b_j\geq0\,,a\geq0 \qquad j=1,2,\ldots,p$$

تعریف ۱–۱۳ مدل اتورگرسیو تعمیم‌یافته مشروط به ناپایداری واریانس

فرآیند $\{r_t\}$ از مدل اتورگرسیو تعمیم‌یافته مشروط به ناپایداری واریانس[3] $GARCH(p,q)$ پیروی می‌کند اگر

[2] Auto Regressive Cinditional Heteroscedasticity
[3] General Auto Regressive Conditional Heteroscedasticity

$$r_t = \sigma_t \varepsilon_t \qquad t = 1,2,\ldots$$

$$\sigma_t^2 = a + \sum_{i=1}^{p} b_i r_{t-i}^2 + \sum_{j=1}^{q} c_j \sigma_{t-j}^2$$

که

$$\{\varepsilon_t\} \sim IID(0,1)$$

و ε_t برای هر t از $\{r_{t-k}, k \geq 1\}$ مستقل می‌باشد.

برای اطمینان از ثابت بودن واریانس شرطی مدل باید قیدهای زیر را در نظر بگیریم:

$$a \geq 0 \ , b_i \geq 0 , \ c_j \geq 0 .$$

این مدل نخستین بار در سال ۱۹۸۶ توسط بولرسلو [7] معرفی شد.

قضیه۱–۱ قضیه مقدار میانگین

هرگاه f یک تابع پیوستهٔ حقیقی بر $[a,\mathrm{b}]$ باشد که در (a,b) مشتقپذیر است، آنگاه نقطه‌ای مانند $x \in (a,b)$ هست که در آن

$$f(b) - f(a) = (b-a) f'(x).$$

(اثبات در [30] صفحه ۱۳۳).

فصل ۲

معرفی مدل توزیع انتقال آمیخته

۲-۱ مقدمه

گاهی اوقات در بررسی داده‌های سری زمانی، با استفاده از ترکیب متناهی از توزیع‌ها می‌توان پدیده‌های تصادفی متنوعی را مدلبندی نمود. کاربرد مدل‌های متناهی آمیخته[4] (MTD) در زمینه‌های مختلفی همچون ستاره‌شناسی، زیست‌شناسی، ژنتیک، داروسازی، اقتصاد، مهندسی و غیره است. بیش از ۱۰۰ سال پیش، یکی از اولین تحلیل‌های آماری با استفاده از مدل‌های آمیخته توسط زیست‌شناس مشهور به نام کارل پیرسن [5] انجام شد. او در سال ۱۸۹۴ آمیخته‌ای از دو تابع چگالی احتمال نرمال را با میانگین‌های متفاوت μ_1 و μ_2 و واریانس‌های σ_1^2 و σ_2^2 با نسبت‌های π_1 و π_2 به داده‌های تهیه شده توسط ولدان[6] (۱۸۹۲و۱۸۹۳) برازش داد. سپس برازش پارامترهای مدل متناهی آمیخته با استفاده از

---

4 Finite Mixture Models
5 Karl Pearson
6 Weldon

الگوریتم EM[۷] (امیدریاضی- ماکسیمم‌سازی) در مقالات مختلفی نظیر والف[۸] (۱۹۶۵) در [25] و همچنین دمپستر، لایر و رابین[۹] (۱۹۷۷) در [10] بحث و مطرح شده است.

در ابتدا تعریف مختصری از مدل‌های آمیخته را بیان و سپس مدل توزیع انتقال آمیخته[۱۰] MTD که اولین بار توسط رفتری در سال ۱۹۸۵ (رجوع کنید به [22]) برای مدلبندی زنجیرهای مارکف مرتبه بالا با فضای وضعیت متناهی را مطرح کرد، معرفی کرده و در ادامه دو تعمیم از این مدل‌ها ارائه می‌شود.

۲-۲ مدل متناهی آمیخته

فرض کنید $Y_1,\ldots,Y_n$ نمونه تصادفی n تایی است به‌طوریکه Y_j یک بردار تصادفی p-بعدی با تابع چگالی احتمال $f(y_j)$ روی فضای R^p باشد. برای مثال اگر $y=\left(y_1^T,\ldots,\mathrm{y}_\mathrm{n}^\mathrm{T}\right)^T$ یک نمونه تصادفی مشاهده‌شده که y_j مقدار مشاهده‌شده بردار تصادفی Y_j باشد، آنگاه چگالی $f(y_j)$ بردار تصادفی Y_j به فرم زیر نوشته می‌شود:

$$f(y_j)=\sum_{i=1}^{g}\pi_i f_i(y_j) \qquad (۱-۲)$$

که $f_i(y_j)$ چگالی‌های مؤلفه‌ای مدل متناهی آمیخته[۱۱] هستند و π_i کمیت‌های غیرمنفی هستند که مجموع آنها برابر یک است به‌طوریکه

$$0\le\pi_i\le 1,(i=1,\ldots,g) \qquad (۲-۲)$$

$$\sum_{i=1}^{g}\pi_i=1 \qquad (۳-۲)$$

---

7 Expectation-Maximization Algorithm
8 Wolfe
9 Dempster, Laird & Rubin

10 Mixture Transition Distribution Model
11 Finite Mixture Model

کمیت‌های $\pi_1,\dots,\pi_g$ وزن‌ها یا کمیت‌های آمیخته نامیده می‌شوند. چگالی (۲-۱)، چگالی آمیخته متناهی g-مؤلفه‌ای و متعاقباً $F(y_j)$ تابع توزیع آمیخته متناهی g-مؤلفه‌ای نامیده می‌شود(برای اطلاعات بیشتر مراجعه کنید به [21]).

۲-۳ مدل توزیع انتقال آمیخته

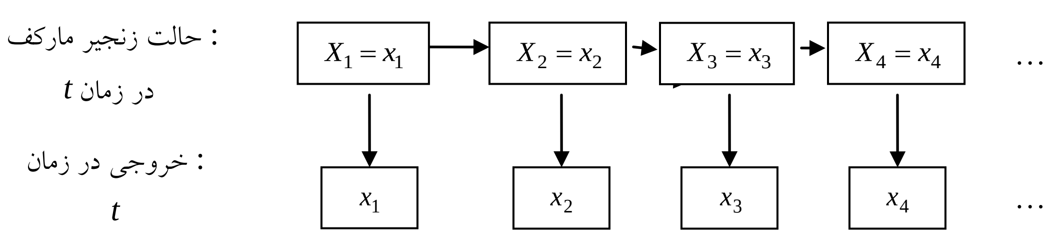
در این بخش به اختصار به یادآوری مفاهیمی از زنجیرهای مارکف پرداخته و سپس مدل توزیع انتقال آمیخته را معرفی و تفاوت آن را با زنجیرهای مارکف بیان می‌کنیم.

2-۳-۱ زنجیر مارکف

بسیاری از دستگاه‌ها دارای این خاصیت‌اند که وقتی وضعیت کنونی آنها معلوم باشد، وضعیت‌های گذشته هیچ اثری بر وضعیت‌های آینده ندارد. این خاصیت به خاصیت مارکوفی موسوم است و دستگاه‌هایی که دارای این خاصیت‌اند، زنجیرهای مارکوفی نامیده می‌شوند. تعاریف و مفاهیم زیر خلاصه‌ای از موارد بیان شده در [3] است (برای اطلاعات بیشتر رجوع کنید به [22] و [3]).

۲-۳-۲ زنجیر مارکف همگن مرتبه اول

$\{X_t\}$ را یک دنباله از متغیرهای تصادفی زمان‌گسسته در نظر گرفته که از مجموعه متناهی $\{1,\dots,m\}$ مقدار می‌پذیرد. در زنجیر مارکف همگن مرتبه اول[12]، فرض می‌شود کل گذشته یک فرآیند تصادفی تنها به‌وسیله آگاهی از آخرین مشاهده خلاصه شود. شکل زیر رابطه وابستگی مشاهدات متوالی در یک زنجیر مارکف را نشان می‌دهد.



[12] First-order homogenous Markov chain