

مدل سازی سیستم های خطی تغییرناپذیر با زمان توسط الگوریتم ژنتیکی پیوسته

سید مصطفی کلامی هریس

دانشکده مهندسی برق و کامپیوتر دانشگاه تبریز

sm.kalami@gmail.com

چکیده - در تحلیل و طراحی سیستم های کنترلی، مدل سازی سیستم به شکل ساده و صحیح، یکی از مهم ترین کارهایی است که می بایست انجام شود. زیرا تمام محاسبات بر اساس صحت مدل پایه ریزی می شود. اما توصیف های تحلیلی برای اکثر سیستم ها یا وجود ندارند و یا در صورت وجود، بسیار پیچیده و عملاً غیر قابل استفاده هستند. برای حل این گونه مسائل، می توان از روش های بهینه سازی هوشمند، همچون الگوریتم های ژنتیکی استفاده کرد. در این مقاله به بحث در مورد الگوریتم ژنتیکی پیوسته ای پرداخته ایم که برای مدل سازی یک سیستم فرضی از روی داده های زمانی نوشته شده است. نتایج حاصل از اجرای این الگوریتم طبقه بندی شده اند و پارامترهای مختلف آن، از نظر اثری که بر پاسخ نهایی دارند مورد بررسی قرار گرفته اند.

کلید واژه ها - الگوریتم های ژنتیکی، پیش بینی سری زمانی، روش های بهینه سازی، شناسایی سیستم، مدل سازی خطی

۱- مقدمه

مقاله برای مدل سازی سیستم های خطی تغییرناپذیر با زمان، از الگوریتم ژنتیکی پیوسته ای استفاده شده است که شرح و توضیح آن در ادامه خواهد آمد.

یک سری زمانی عبارت است از رشته ای از مقادیر اندازه گیری شده از متغیرهای حالت سیستم در یک بازه زمانی محدود. در اغلب اوقات منظور ما از مدل کردن سیستم ها، پیش بینی کردن مقادیر سری زمانی در آینده است [۵]. نتایج این پیش بینی ها در زمینه های مختلف قابل استفاده هستند. به عنوان نمونه اقتصاد، هواشناسی و مقاصد مهندسی همچون پیش بینی بار خطوط نیرو [۲]، شبیه سازی سیستم ها و بررسی رفتار سازه ها از زمینه هایی هستند که پیش بینی سری زمانی در آن ها کاربرد فراوان دارد [۵].

این مقاله مشتمل بر این بخش ها می باشد. تعریف مسئله مورد بررسی، به همراه توضیح مختصری درباره الگوریتم به کار رفته برای حل مسئله، در بخش ۲ آمده است. الگوریتم با تنظیمات مختلف اجرا شده است و نتایج حاصل از این

یکی از مهم ترین ارکان علم کنترل، توصیف پدیده های فیزیکی به صورت روابط ریاضی است، به نحوی که بتوان ایده های کنترلی را روی آن ها پیاده کرد. در تحلیل و طراحی سیستم های کنترلی، مدل سازی به شکل ساده و صحیح، یکی از مهم ترین کارهایی است که می بایست انجام شود. زیرا تمام محاسبات بر اساس صحت مدل پایه ریزی می شود. اما توصیف های تحلیلی برای اکثر سیستم ها یا وجود ندارند و یا در صورت وجود، بسیار پیچیده و عملاً غیر قابل استفاده هستند. به خصوص در مواردی که یک یا چند مورد از متغیرهای موجود در سیستم، از رابطه یا روابط خاص ریاضی تبعیت نکنند، نمی توان سیستم را به صورت بسته و توسط یک رابطه ریاضی توصیف کرد و مسئله به صورت تحلیلی قابل حل نخواهد بود.

برای حل مسائلی که توسط روش های سنتی قابل حل و بررسی نمی باشند، می توان از روش های بهینه سازی هوشمند همچون روش الگوریتم های ژنتیکی بهره برد [۱]. در این

بیشتر در پیش‌بینی عملکرد سیستم‌ها می‌بایست اطلاعاتی با حجم نسبتاً زیاد، یعنی با K زیاد و همچنین با محتوای بالا، یعنی با T_s کم از سیستم داشته باشیم.

مسئله مورد بررسی یافتن ماتریس حالت A است به نحوی که بتوانیم مقادیر سری زمانی زیر را پیروی کنیم :

$$\mathbf{x}_s[k] = \begin{bmatrix} \sin k.T_s \\ \cos k.T_s \end{bmatrix} \quad (4)$$

$$k = 0, 1, \dots, 500$$

$$T_s = \frac{\pi}{50}$$

البته به دلیل وجود خطا در محاسبه مشتق عددی [۴]، رسیدن به هزینه صفر غیر ممکن است. یعنی کمترین هزینه قابل دسترسی یک مقدار غیر صفر می‌باشد که در واقع برای شرایط مسئله ما، یک مینیمم مطلق محسوب می‌شود.

در الگوریتم ژنتیکی مورد بحث، درایه‌های ماتریس A به عنوان ژن‌های کروموزوم در نظر گرفته شده‌اند. همچنین ترکیب خطی ژن‌ها با یک ضریب تصادفی به عنوان عمل تقاطع یا Cross-over مورد استفاده قرار گرفته است. در هنگام عمل جهش یا Mutation نیز فقط یک ژن از ژن‌های کروموزوم را با یک عدد تصادفی تعویض می‌کنیم.

برای پایان‌پذیری الگوریتم نیز از سه روش متداول [۳ و ۱] استفاده شده است که عبارتند از : کم‌تر شدن مقدار تابع هزینه از یک مقدار مطلوب ، ثابت ماندن مقدار تابع هزینه برای تعداد معینی از نسل‌ها و پشت سر گذاشته شدن تعداد معینی از نسل‌ها.

۳- نتایج به دست آمده از اجرای الگوریتم

نمودار بهترین (کم‌ترین) مقدار تابع هزینه و هزینه متوسط الگوریتم ژنتیکی مورد بحث برای یک مورد اجرا، در شکل ۱ آمده است. می‌بینیم که الگوریتم تقریباً در نسل بیست و دوم جواب بهینه را یافته است که همان مینیمم مطلق قابل دستیابی برای الگوریتم است. همچنین روند کاهشی مقدار تابع هزینه در طول اجرا نشان دهنده همگرایی سریع الگوریتم می‌باشد.

اجراها در بخش ۳ آمده است. بخش ۴ نیز شامل نتیجه‌گیری‌های نهایی در خصوص الگوریتم مورد بحث می‌باشد.

۲- طرح مسئله و الگوریتم به کار رفته

پس از آزمایش روی سیستمی، متغیرهای حالت مختلف آن را در لحظات مختلف از زمان و با فاصله زمانی معین T_s اندازه‌گیری کرده‌ایم. سری زمانی حاصل از این اندازه‌گیری‌ها تابع برداری گسسته در زمانی به شکل $\mathbf{x}_s[k]$ است، که اصطلاحاً آن را سری زمانی می‌نامیم. به کمک روش‌های عددی، مشتق تقریبی [۴] این سری زمانی را نسبت به زمان محاسبه می‌کنیم و به صورت $\dot{\mathbf{x}}_s[k]$ نشان می‌دهیم. می‌خواهیم ماتریس حالت A را پیدا کنیم به نحوی که به ازای همه مقادیر مجاز k داشته باشیم :

$$\dot{\mathbf{x}}_s[k] = A.\mathbf{x}_s[k] \quad (1)$$

هدف از حل چنین معادله‌ای یافتن یک مدل پیوسته در زمان برای سیستم مورد مطالعه است. اگر این معادله حل شود، می‌توان رفتار سیستم را در آینده پیش‌بینی کرد و یا توسط ایده‌های کنترلی عملکرد آن را بهینه نمود.

برای حل این مسئله توسط الگوریتم ژنتیکی پیوسته ابتدا می‌بایست تابع هزینه مناسبی را طرح نمود [۱]. بردار خطا را که عبارت است از اختلاف مشتق تقریبی و مشتق واقعی رشته $\mathbf{x}_s[k]$ ، به شکل زیر تعریف می‌کنیم:

$$\mathbf{e}[k] = \dot{\mathbf{x}}_s[k] - A.\mathbf{x}_s[k] \quad (2)$$

تعداد نمونه‌های موجود در $\mathbf{x}_s[k]$ را K و بعد سیستم را n در نظر می‌گیریم، و مجموع مقادیر نرم مرتبه p بردارهای خطا را به عنوان تابع هزینه تعریف می‌کنیم. اگر منظور از $e_{k,i}$ درایه i ام بردار $\mathbf{e}[k]$ باشد ، تعریف تابع هزینه به شکل زیر خواهد بود:

$$\text{cost} = \sum_{k=0}^{K-1} \|\mathbf{e}[k]\|_p = \sum_{k=0}^{K-1} p \sqrt[p]{\sum_{i=1}^n e_{k,i}^p} \quad (3)$$

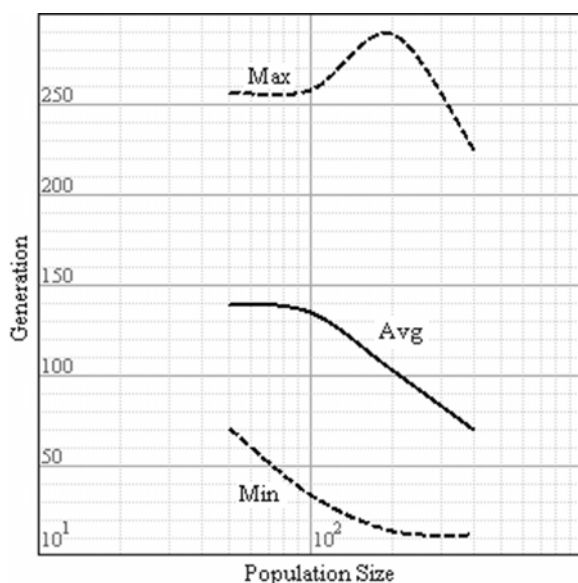
واضح است که مقادیر T_s و K در کیفیت پاسخ به دست آمده بسیار موثر خواهند بود. برای داشتن دقت و صحت

مینیمم مطلق برای شرایط فعلی مسئله است. مقدار کمینه ۰/۰۶۲۸۱۱ برای تابع هزینه، ناشی از خطای مشتق گیری عددی است و با کمی دقت در می یابیم که :

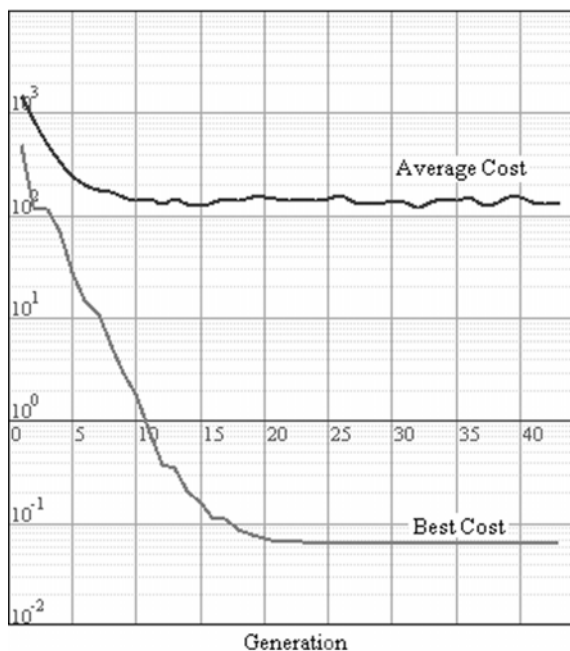
$$T_s = \frac{\pi}{50} \approx 0.062811 = \text{cost}_{\min} \quad (5)$$

در شکل ۲ نمودار تعداد نسل های طی شده برای رسیدن به جواب بر حسب جمعیت آمده است. در شکل ۳ نیز نمودار متوسط، بیشینه و کمینه هزینه بر حسب اندازه جمعیت ترسیم شده است. مثلاً متوسط هزینه برای اجرایی که جمعیت در آن ها برابر با ۱۰۰ بوده است، تقریباً برابر با ۲ می باشد.

شکل های ۲ و ۳ هر دو نشان دهنده میزان تاثیر قابل توجه اندازه جمعیت بر سرعت و دقت الگوریتم هستند. سرعت الگوریتم برای جمعیت های کمتر از ۱۰۰ تقریباً در حد ثابتی است، اما با افزایش جمعیت، به تدریج سرعت رسیدن به جواب نیز بالاتر می رود. از نظر مقدار تابع هزینه جمعیت بیشتر همواره موجب کمتر شدن مقدار تابع هزینه می شود. اما چون این مقدار هیچ گاه صفر نمی شود، افزایش بیش از حد جمعیت سود چندانی ندارد و فقط باعث کندتر شدن روند اجرای الگوریتم می شود. جمعیت ۴۰۰ برای این الگوریتم آستانه اشباع است. یعنی اجرای الگوریتم برای جمعیت های بیشتر از ۴۰۰ تفاوت زیادی در نتیجه به وجود نخواهد آورد.



شکل ۲: نمودار تعداد نسل ها برای رسیدن به جواب بر حسب جمعیت



شکل ۱: هزینه بهینه و هزینه متوسط یک اجرای نمونه

پارامترهای الگوریتم ژنتیکی [۱] که در زمان اجرا تغییر داده شده اند، جمعیت هر نسل N_{pop} ، درصد ماندگاری هر نسل برای نسل بعدی X_{rate} ، درصد جهش μ و درجه نرم مورد استفاده p هستند. مقادیر مختلف که برای این پارامترها اختیار شده اند، عبارتند از :

$$N_{pop} = 50, 100, 200, 400$$

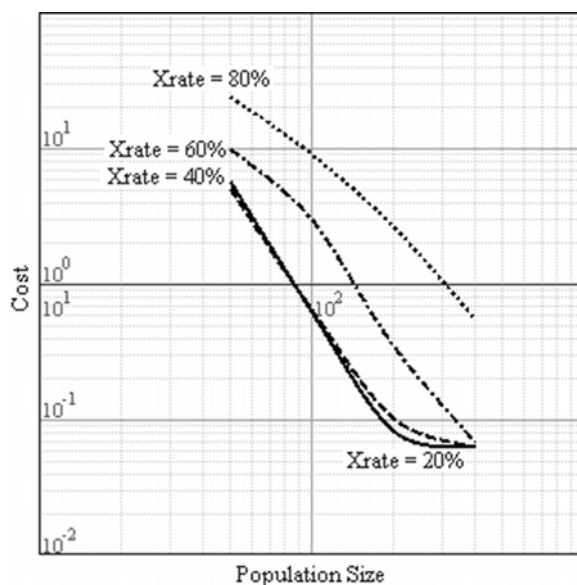
$$X_{rate} = 20\%, 40\%, 60\%, 80\%$$

$$\mu = 0\%, 5\%, 10\%, 20\%, 50\%, 75\%$$

$$p = 1, 2, 4, 10, 100$$

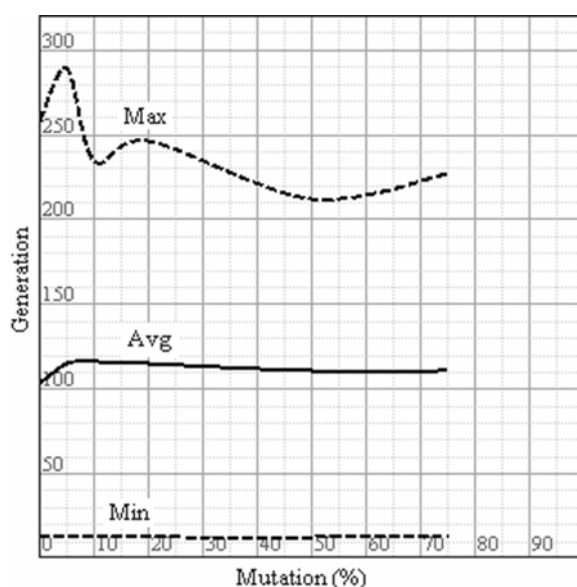
لذا ۴۸۰ ترکیب مختلف از این مقادیر قابل آزمایش می باشند. هر کدام از این موارد ۵۰ بار اجرا شده اند و نتایج حاصل از این ۲۴۰۰۰ بار اجرا، جمع آوری شده اند که در ادامه مورد بحث و بررسی قرار خواهند گرفت.

این الگوریتم به طور متوسط در ۱۱۲/۱ نسل جواب را یافته است و مقدار متوسط تابع هزینه طی موارد اجرا ۰/۸۶۵۲۶۱ بوده است. الگوریتم در سریع ترین حالت در ۱۲/۳ نسل و در کندترین حالت در ۲۸۹/۳ نسل به جواب رسیده است. بهترین مقدار تابع هزینه ۰/۰۶۲۸۱۱ و بدترین مقدار ۱۲۴/۶۴۹۹۰۶ بوده است. همان طور که می بینیم حتی در بهترین شرایط، باز هم مقدار تابع هزینه صفر نمی شود. این مقدار کمترین مقدار هزینه ای است که با اطلاعات موجود در سری زمانی می توان به دست آورد، و به اصطلاح یک

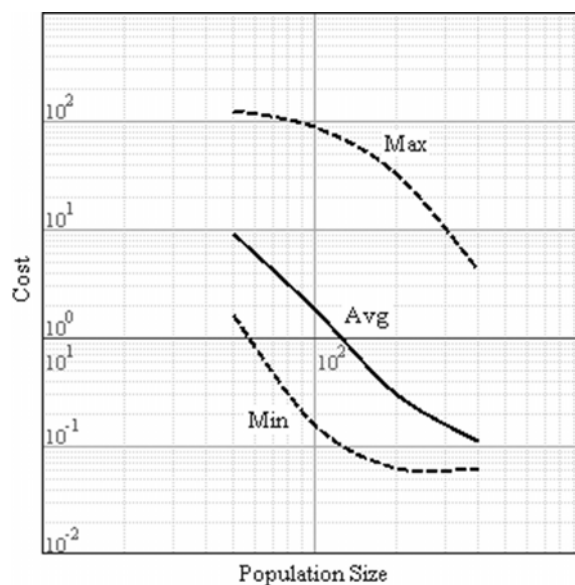


شکل ۵: نمودار متوسط مقدار تابع هزینه بر حسب جمعیت که بر اساس مقدار درصد ماندگاری طبقه‌بندی شده است.

در شکل های ۴ و ۵ به وضوح دیده می شود که X_{rate} زیاد اصلا به نفع ما نیست. اگر X_{rate} را کاهش دهیم هم سرعت الگوریتم افزایش می یابد و هم هزینه کمتر می شود. اصولا رسیدن به بهترین جواب با همه مقادیر X_{rate} امکان پذیر است. اما قطعا با افزایش X_{rate} سرعت رسیدن به جواب کاهش چشمگیری خواهد یافت. مقدار بهینه برای X_{rate} بین ۲۰٪ و ۴۰٪ می باشد. اثر X_{rate} به اندازه اثر جمعیت مهم است و نمی توان گفت کدام یک مهم تر یا موثرترند. ولی به هر حال اثر این دو عامل بیشتر از بقیه است.

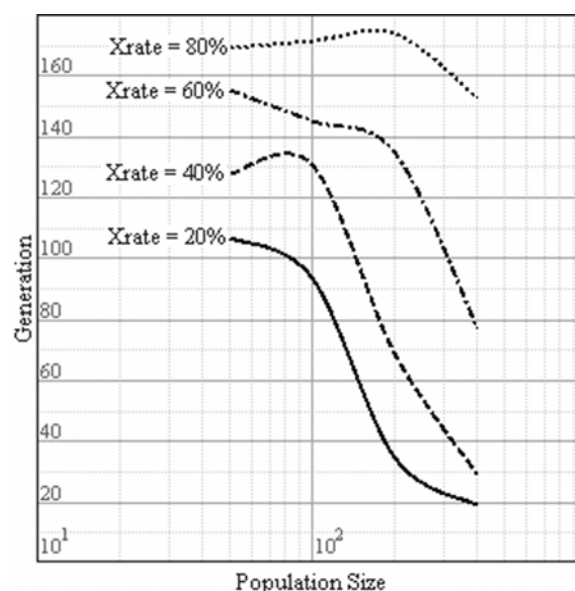


شکل ۶: کمینه، متوسط و بیشینه تعداد نسل ها بر حسب جهش



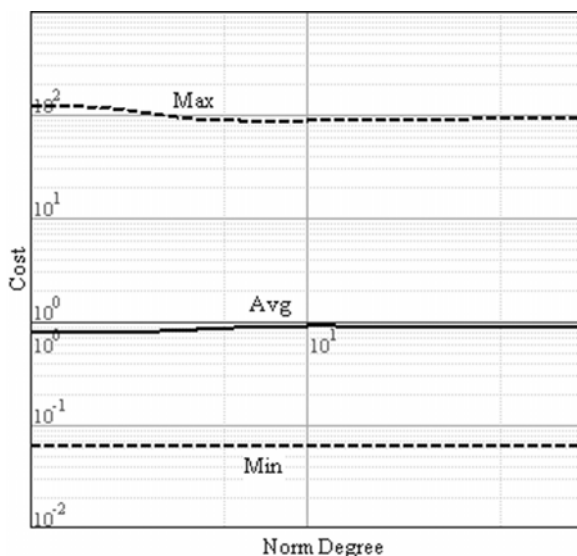
شکل ۳: نمودار کمینه، متوسط و بیشینه هزینه بر حسب جمعیت

در شکل های ۴ و ۵ نیز به ترتیب نمودار متوسط تعداد نسل ها برای رسیدن به جواب و متوسط هزینه بر حسب جمعیت ترسیم شده اند. تفاوت این نمودارها با نمودارهای قبلی، در طبقه‌بندی شدن نمودارها بر حسب درصد ماندگاری X_{rate} می باشد. لذا در هر شکل ۴ نمودار مختلف متناظر با ۴ مقدار مختلف درصد ماندگاری می بینیم. این نوع طبقه‌بندی بررسی تاثیر دو عامل جمعیت و درصد ماندگاری را به صورت هم زمان ممکن می کند.



شکل ۴: نمودار متوسط تعداد نسل بر حسب جمعیت که بر اساس مقدار درصد ماندگاری طبقه‌بندی شده است.

عامل چهارم که درجه نرُم مورد استفاده در محاسبه تابع هزینه می‌باشد، تقریباً بی‌تاثیر است. اما با افزایش درجه نرم سرعت الگوریتم با افزایش نسبی همراه است. به خصوص در جمعیت‌های بزرگ این افزایش قابل مشاهده‌تر خواهد بود. اما قطعاً درجه نرم مورد استفاده روی مقدار متوسط تابع هزینه بی‌تاثیر است. در شکل‌های ۸ و ۹ نمودار تعداد نسل‌ها و مقدار تابع هزینه بر حسب درجه نرم دیده می‌شوند.



شکل ۹: نمودار کمینه، متوسط و بیشینه مقدار تابع هزینه بر حسب درجه نرُم مورد استفاده

حاصل بررسی‌ها نشان می‌دهد که، می‌توان تنظیماتی برای الگوریتم انتخاب کرد که کمترین بار محاسباتی و به همراه آن بهترین نتیجه را به دست آورد. این ترکیب از تنظیمات را ترکیب بهینه می‌نامیم. ترکیب بهینه به دست آمده از آزمایش‌ها، که بهترین بازده داشته است، عبارت است از:

$$N_{pop} = 200$$

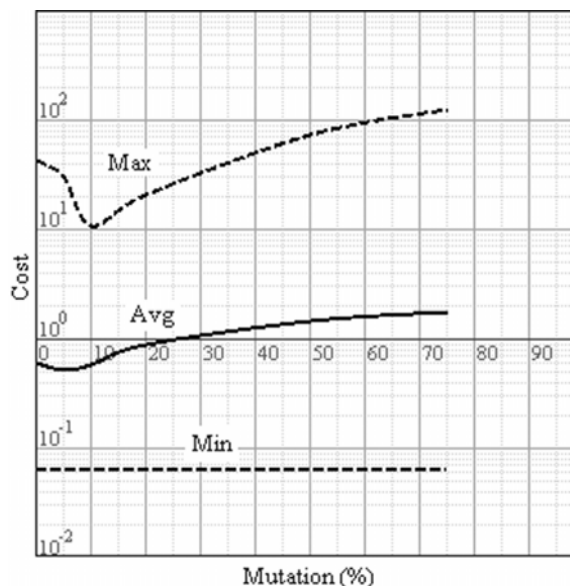
$$X_{rate} = 20\%$$

$$\mu = 5\%$$

$$p = 100$$

۴- نتیجه‌گیری

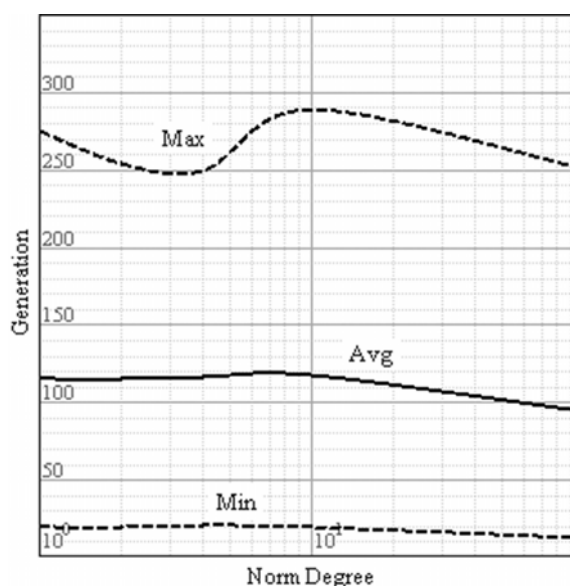
در این مقاله برای مدل‌سازی یک سیستم خطی تغییرناپذیر با زمان، از یک الگوریتم ژنتیکی پیوسته استفاده شده است. این کار که به نام پیش‌بینی سری زمانی توسط الگوریتم‌های ژنتیکی مشهور است، منجر به ارائه یک مدل خطی با مرتبه دلخواه از روی داده‌های زمانی هر سیستمی می‌شود. نتایج



شکل ۷: نمودار کمینه، متوسط و بیشینه هزینه بر حسب جهش

در شکل‌های ۶ و ۷ نیز نمودارهای تعداد نسل برای رسیدن به جواب و مقدار تابع هزینه بر حسب درصد جهش آمده‌اند. این نمودارها نشان می‌دهند که عامل جهش تاثیر چندانی بر سرعت یا دقت الگوریتم ندارد.

اما به هر حال مقادیر بین ۰٪ و ۲۰٪ بهترین مقادیری هستند که می‌توان برای میزان جهش در نظر گرفت. به ازای همه مقادیر جهش الگوریتم ژنتیکی به صورت بالقوه می‌تواند به بهترین جواب ممکن برسد و بودن یا نبودن جهش تقریباً بی‌تاثیر است.



شکل ۸: نمودار کمینه، متوسط و بیشینه تعداد نسل‌ها برای رسیدن به جواب بر حسب درجه نرُم مورد استفاده

به دست آمده از اجرای الگوریتم ژنتیکی، حاکی از آن است که اندازه جمعیت و درصد ماندگاری، بیشترین تاثیر را در پاسخ الگوریتم دارند و سایر عوامل تاثیر بسزایی در سرعت یا دقت الگوریتم ندارند.

سپاسگزاری

از زحمات و راهنمایی‌های بی‌دریغ استاد گرانقدر جناب آقای دکتر محمدتقی وکیلی، نهایت تشکر و امتنان را دارم.

مراجع

- [1] R. L. Haupt and S. E. Haupt, *Practical Genetic Algorithms*, John Wiley & Sons Inc. Publication, 2004
- [2] V. Varadan, H. Leung and É. Bossé, "Dynamical Model Reconstruction and Accurate Prediction of Power-Pool Time Series" IEEE Trans. Instrumentation, Vol. 55, No. 1, pp. 327-336, Feb 2006
- [3] M. Melanie, *An Introduction to Genetic Algorithms*, MIT Press, 1999
- [4] R. J. Schilling and S. L. Harris, *Applied Numerical Methods for Engineers*, Brooks/Cole, 2000
- [5] D. Kim and C. Kim, "Forecasting Time Series with Genetic Fuzzy Predictor Ensemble" IEEE Trans. Fuzzy Systems, Vol. 5, No. 4, pp. 523-535, Nov 1997