

مقدمه‌ای بر داده‌های دایره‌ای

شیمای حاجی‌زاده^۱، مجید سرمد^۲

چکیده:

در بسیاری از زمینه‌های مختلف علمی، اندازه‌گیری‌ها جهت‌دار هستند. برای مثال، یک زیست‌شناس ممکن است جهت پرواز یک پرنده یا جهت حرکت یک حیوان را اندازه‌گیری کند. در این مقاله، ابتدا داده‌های دایره‌ای معرفی نشده است و سپس روش محاسبه میانگین جهت و همچنین پراکندگی و گشتاورهای بالاتر مربوط به داده‌های دایره‌ای ارائه شده‌اند. بسیاری از مسائل داده‌های دایره‌ای غالباً در فرم‌های تحلیلی تقریباً ساده قابل حصول نیستند، نرم‌افزارهای کامپیوتری مرتبط با استفاده از این روش‌ها ضروری است؛ که در پایان این نوشتار، بسته توابع CircStats در نرم‌افزار R و Matlab برای تحلیل مجموعه داده‌های جهت‌دار ارائه شده است.

واژه‌های کلیدی: داده‌های دایره‌ای، فاصله دایره‌ای، میانگین جهت، Matlab، R.

۱ مقدمه

سه بعد ممکن است با دو زاویه (شبیبه به نمایش نقاط روی سطح کره زمین با طول و عرض جغرافیایی آن‌ها) به صورت بردارهای واحد در سه جهت یا به صورت نقاط روی سطح یک کره واحد نشان داده شوند. به این دلیل داده‌های جهت‌دار در سه بُعد، نیز داده‌های کروی^۴ نامیده می‌شوند. داده‌های جهت‌دار دارای ویژگی‌های منحصر به فردی هستند. برای مثال نمایش عددی یک جهت دو بعدی به صورت زاویه یا یک بردار واحد لزوماً منحصر به فرد از مقدار زاویه‌ای وابسته به انتخاب چیزی که به عنوان جهت صفر و جهت چرخش علامت‌دار شده است، نیست. آنچه که توسط یک ریاضیدان 60° در نظر گرفته شده است، به گونه‌ای است که شرق حقیقی را به عنوان جهت صفر و خلاف جهت عقربه‌های ساعت را به عنوان جهت مثبت اختیار کرده است. در حالی که همان جهت، توسط یک زمین‌شناس که شمال حقیقی را به عنوان جهت صفر و حرکت در جهت عقربه‌های ساعت را به عنوان جهت مثبت اختیار کرده است، 30° در نظر گرفته می‌شود، که «azimuth»^۵ نامیده می‌شود.

به طور مشابه، مطلب فوق برای مقادیر اختصاص داده شده به هر مجموعه داده کروی نیز درست است، لذا مهم است تا مطمئن شویم که نتایج (آمار توصیفی و استنباطی) براساس مشاهدات

در بسیاری از زمینه‌های مختلف علمی، اندازه‌گیری‌ها جهت‌دار هستند. برای مثال یک زیست‌شناس ممکن است جهت پرواز یک پرنده یا جهت حرکت یک حیوان را اندازه‌گیری کند، در حالی که یک زمین‌شناس ممکن است علاقه‌مند به جهت قطب مغناطیسی زمین باشد. جهات، ممکن است در دو بُعد مانند جهت پرواز یک پرنده یا سه بُعد همچون جهت قطب مغناطیسی زمین باشد. یک مجموعه از چنین مشاهداتی روی جهات به عنوان داده‌های جهت‌دار منسوب می‌شوند. جهت‌های دو بعدی را می‌توان به صورت زوایای اندازه‌گیری شده با توجه به انتخاب مناسب «جهت صفر»، به عنوان نقطه شروع و «جهت چرخش»، چه در جهت عقربه‌های ساعت و یا خلاف جهت عقربه‌های ساعت به عنوان جهت مثبت اختیار کرد. از آنجایی که یک جهت دارای مقدار و اندازه نیست، این‌ها را می‌توان به راحتی به صورت نقاطی روی محیط یک دایره واحد متمرکز در مبدأ یا به صورت بردارهای واحد که مبدأ را به این نقاط متصل می‌کند، نشان داد. به دلیل این نمایش دایره‌ای، مشاهدات روی این قبیل جهات دو بعدی داده‌های دایره‌ای^۳ نیز نامیده می‌شوند. به طور مشابه، جهات در

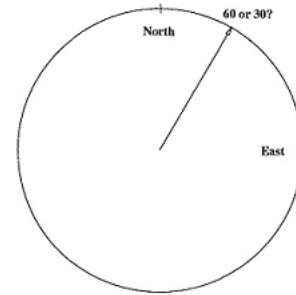
^۱ کارشناس ارشد آمار اقتصادی و اجتماعی، دانشگاه فردوسی مشهد

^۲ استادیار گروه آمار، دانشگاه فردوسی مشهد

^۳ Circular data

^۴ Spherical data

داده شده هستند و وابسته به مقادیر دلخواه نیستند. بنابراین باید به نتایجی برسیم که وابسته به انتخاب دلخواه و جهت چرخش نیستند. در این مقاله، جزئیات بسته^۶ *CircStats* در نرم‌افزار R و Matlab برای تحلیل مجموعه داده‌های جهت‌دار بیان شده است.



شکل ۱. داده‌های جهت‌دار به انتخاب مبدأ و جهت چرخش وابسته است.

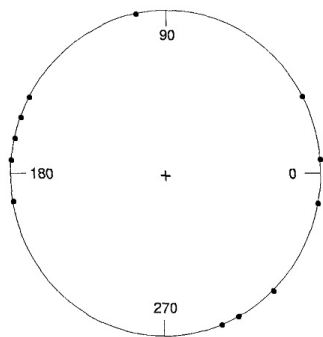
۲ کاربردها و زمینه‌ها

داده‌های جهت‌دار در دو یا سه بُعد، به طور مکرر در علوم فیزیکی و طبیعی چون زیست‌شناسی، پزشکی، محیط‌زیست، زمین‌شناسی و... به کار می‌روند. به طور مثال، زیست‌شناسان در مطالعه مهاجرت پرندگان، جهت پرواز پرندگان را که به تنهایی رها می‌شوند، ثبت می‌کنند. در حالی که برخی دیگر، حرکات یا جهتگیری‌های برخی از موجودات بیولوژیکی را ثبت می‌کنند. در یک آزمایش روی کبوترهای خانگی اسمیت - کوینگ^۷، مجموعه داده‌های زیر براساس زوایای ناپدید شدن ۱۵ گونه پرنده که به تنهایی رها شده‌اند، به دست آمده‌اند (به صورت افزایشی مرتب شده‌اند)

85, 135, 135, 140, 145, 150, 150,

160, 185, 200, 210, 220, 225, 270.

داده‌ها به صورت نقاطی روی دایره در شکل ۲ نشان داده شده‌اند. جدول ۱ داده‌های مربوط به جهت‌گیری لاک‌پشت‌ها بعد از تخم‌گذاری را نشان می‌دهد.



شکل ۲. زوایای ناپدید شدن کبوترهای رها شده در آزمایش

اسمیت - کوینگ

^۶Package

^۷Schmidt-Koeing

ترتیب یا رتبه بندی طبیعی از مشاهدات وجود ندارد، از این رو، این که آیا یک جهت «بزرگتر» از یک جهت دیگر است یا نه، بستگی دارد که آیا در جهت عقربه‌های ساعت یا خلاف جهت عقربه‌های ساعت به عنوان جهت مثبت رفتار کرده است یا نه و نیز کجا «صفر» است. این باعث می‌شود که روش‌های مبتنی بر رتبه اساساً غیر قابل اجرا باشد.

از آنجاکه تحلیل داده‌های جهت‌دار تفاوت قابل ملاحظه‌ای با تحلیل داده‌های خطی دارند، روش‌ها و اندازه‌گیری‌های آماری با توجه به انتخاب جهت صفر و جهت چرخش دلخواه، باعث می‌شود که بسیاری از روش‌ها و اندازه‌گیری‌های خطی معمول اغلب همراه‌کننده باشند. بنابراین بسیاری از مفاهیم چون میانگین، واریانس، تابع مولد گشتاور، همبستگی، رگرسیون و... نیاز به ابداع برای داده‌های دایره‌ای دارند. به همین ترتیب، ایده‌های استنباط آماری همچون ناریبی، توابع زیان، کران‌های واریانس، یکنوایی توابع توان از آزمون‌ها و... نیاز دارند که مجدداً با احتیاط و با ملاحظات لازم تعریف شوند. ماردیا [۶]، فیشر [۲]، جامالاماداکا و سارما [۴]، جامالاماداکا و سنگوپتا [۵] و بسیاری از افراد دیگر به تحقیق درباره داده‌های جهت‌دار (دایره‌ای) در زمینه آمار استنباط و مباحث مربوط به همبستگی بین متغیرها، رگرسیون و... پرداخته‌اند.

از آنجا که راه‌حل‌های بسیاری از مسائل داده‌های دایره‌ای، غالباً

جدول ۱: جهت گیری ۷۶ لاک پشت بعد از تخم گذاری

زوایای اندازه گیری شده در جهت عقربه های ساعت از شمال										
۸	۹	۱۳	۱۳	۱۴	۱۸	۲۲	۲۷	۳۰	۳۴	
۳۸	۳۸	۴۰	۴۴	۴۵	۴۷	۴۸	۴۸	۴۸	۴۸	
۵۰	۵۳	۵۶	۵۷	۵۸	۵۸	۶۱	۶۳	۶۴	۶۴	
۶۴	۶۵	۶۵	۶۸	۷۰	۷۳	۷۸	۷۸	۷۸	۸۳	
۸۳	۸۸	۸۸	۸۸	۹۰	۹۲	۹۲	۹۳	۹۵	۹۶	
۹۸	۱۰۰	۱۰۳	۱۰۶	۱۱۳	۱۱۸	۱۳۸	۱۵۳	۱۵۳	۱۵۵	
۲۰۴	۲۱۵	۲۲۳	۲۲۶	۲۳۷	۲۳۸	۲۴۳	۲۴۴	۲۵۰	۲۵۱	
۲۵۷	۲۶۸	۲۸۵	۳۱۹	۳۴۳	۳۵۰					

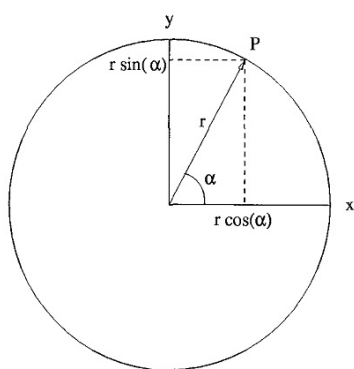
(x, y) براساس مختصات دکارتی یا به صورت (r, α) براساس مختصات قطبی نشان داد، که r فاصله تا مبدأ و برابر یک در نظر گرفته می شود و α جهت آن است. برای نقطه خاص، یعنی مبدأ O ، $r = 0$ و بدون جهت نشان داده شده است (یعنی α تعریف نشده است). همان طور که می دانیم تبدیل مختصات قطبی به مختصات دکارتی با استفاده از توابع مثلثاتی $\sin$ و $\cos$ انجام می شود.

یک نقطه مانند P را با مختصات قطبی (r, α) در نظر بگیرید. از شکل ۳، مختصات دکارتی نقطه P به صورت زیر به دست می آید:

$$x = r \cdot \cos(\alpha), \quad y = r \cdot \sin(\alpha) \quad (۱)$$

همان طور که گفته شد، در تحلیل داده های جهت دار توجه ما به جهت است نه به اندازه بردار، بنابراین ما این بردارها را برای سادگی، به طول واحد (یعنی $r = 1$) در نظر می گیریم. لذا هر جهت، متناظر با نقطه ای روی محیط دایره است. به عبارت دیگر، این نقطه روی محیط یک دایره واحد را می توان فقط با زاویه مشخص کرد. حال اگر نقطه P روی محیط دایره واحد قرار بگیرد، تبدیل بین مختصات قطبی و دکارتی ساده تر خواهد بود:

$$(1, \alpha) \iff (x = \cos(\alpha), y = \sin(\alpha)) \quad (۲)$$



شکل ۳. رابطه بین مختصات قطبی و دکارتی

تذکر ۱.۳. با این که می نویسیم: $\{u_i\} = \{(x_i, y_i)\}$ چون x_i و y_i از طریق معادله $x_i^2 + y_i^2 = 1, \forall i$ مرتبط هستند، چنین داده هایی را نمی توان به صورت داده های دو متغیره روی صفحه در نظر گرفت.

تذکر ۲.۳. به طور مشابه جهت های سه بُعدی را نیز می توان به چند روش معادل نشان داد:

(الف) به صورت نقاطی روی سطح یک کره واحد متمرکز در مبدأ،

جامالاماداکا^۳] و همکاران یک کاربرد جالب توجه در پزشکی که در آن زاویه خم شدن زانو برای ارزیابی بهبود بیماران ارتوپدی اندازه گیری شده بود، مطرح کردند. در محیط زیست جهت باد را به عنوان یک عامل مهم در بسیاری از مطالعات از جمله مطالعاتی که حامل آلاینده ها هستند، در نظر می گیرند.

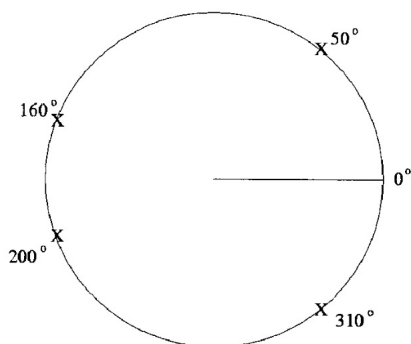
همچنین هر پدیده دوره ای با یک دوره معلوم که یک روز، یک ماه یا یک سال می گویند، می تواند روی دایره جایی که محیط متناظر با این دوره است، واقع شود. به عبارت دیگر در مطالعات شبانه روزی (سیکل) استفاده شود و زمان یک رویداد در این سیکل واقع شود. به عنوان مثال، زمانی که دمای بدن یا فشار خون در طول یک روز اوج می گیرد یا زمان های ورود بیماران با حملات قلبی در یک بیمارستان در طول یک روز. دایره ممکن است برای نشان دادن ۳۶۵ روز سال در نظر گرفته شود و وقوع حوادث هواپیما را روی آن ثبت کنند تا ببینند که آیا آن ها به طور یکنواخت طی فصل های مختلف سال توزیع شده اند یا نه.

۳ آمار توصیفی

همان طور که قبلاً گفته شد، داده های جهت دار را می توان به صورت زوایای یا نقاطی روی محیط دایره واحد ($r = 1$) نشان داد. مکان هر جهت بر روی دایره، می تواند به طور منحصر به فرد (با دو مختصات) تعیین شود. برای این منظور می توانیم دستگاه مختصات دکارتی با مبدأ O و دو محور عمود برهم X و Y از طریق O را استفاده کنیم. هر نقطه مانند P را روی صفحه می توان به صورت

^۳Jammalamadaka

50° ، 160° ، 200° و 300° به روش معمول اندازه‌گیری شده باشند (جهت صفر در شرق و جهت چرخش در خلاف جهت عقربه‌های ساعت به عنوان جهت مثبت در شکل ۵ نشان داده شده‌اند). با اشاره به این که میانگین جهت نقطه‌ای در غرب روی محیط دایره است، میانگین حسابی آنها 180° می‌باشد.



شکل ۵. مقدار میانگین نمونه وابسته به انتخاب مبدأ

از طرف دیگر اگر همان چهار جهت به صورت زوایایی در محدوده -180° تا 180° اندازه‌گیری شوند، با همان جهت صفر و جهت چرخش، مقادیر آنها -50° ، -160° ، 50° و 160° می‌شود. حال میانگین حسابی آنها 0° است، یعنی نقطه‌ای در شرق، که مخالف با میانگین قبلی است. با فرض بعدی که شمال، جهت صفر و جهت چرخش در جهت عقربه‌های ساعت به عنوان جهت مثبت باشد، همان چهار جهت مقادیر 40° ، 140° ، 250° و 290° را می‌گیرند که میانگین حسابی آنها 180° است و مکان آن روی دایره جنوب حقیقی را نشان می‌دهد.

در نتیجه اندازه‌گیری همان چهار جهت در محدوده -180° تا 180° ، مجدداً با شمال به عنوان جهت صفر و چرخش در جهت عقربه‌های ساعت به عنوان جهت مثبت، میانگین حسابی ناشی از شمال را نتیجه می‌دهد. لذا، میانگین حسابی معمول می‌تواند نقطه‌ای در شرق، غرب، شمال یا جنوب (یا نقطه‌ای در بین) برای همان داده‌ها، بسته به انتخاب مبدأ و جهت چرخش باشد.

از مثال‌های فوق، واضح است که چون میانگین حسابی، که معمولاً برای داده‌های خطی استفاده می‌شود، به انتخاب جهت صفر و جهت چرخش وابسته است به عنوان یک معیار تمرکز برای مجموعه داده شده از جهات مناسب نیست. از این رو، باید از آن به عنوان یک معیار تمرکز برای جهات اجتناب شود.

به صورت یک بردار (X, Y, Z) به طول یک.

ب) در شرایطی از دو زاویه (α_1, α_2) در دامنه $[0, \pi] \times [0, 2\pi]$ در شرایط جغرافیایی، این دو زاویه طول و عرض جغرافیایی نامیده می‌شوند. به علاوه می‌توان مختصات قطبی $(1, \alpha_1, \alpha_2)$ را با استفاده از رابطه زیر به مختصات دکارتی تبدیل کرد.

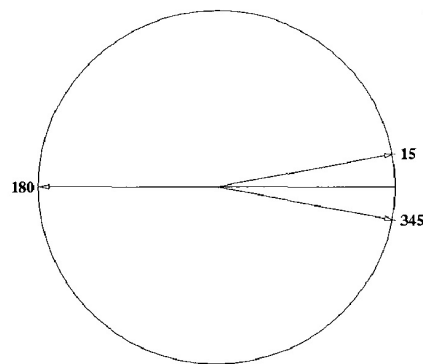
$$(1, \alpha_1, \alpha_2) \iff (x = \sin(\alpha_1) \cos(\alpha_2),$$

$$y = \sin(\alpha_1) \sin(\alpha_2), z = \cos(\alpha_1))$$

۴ معیار تمرکز

به منظور تعریف میانگین جهت که گاهی اوقات «جهت برتر» نیز نامیده می‌شود، با توجه به مجموعه داده‌های جهت‌دار تک‌مُدی^۹ ممکن است برای محاسبه میانگین از میانگین حسابی استفاده کنیم. همان‌طور که قبلاً ذکر شد، میانگین نمونه و همچنین انحراف استاندارد و گشتاورهای بالاتر دیگر، به انتخاب جهت صفر و جهت چرخش وابستگی قوی دارند. دو مثالی که در زیر آمده است، نشان می‌دهد که چرا معیارهای تمرکز فوق به عنوان اندازه‌گیری توصیفی از مرکز برای داده‌های دایره‌ای نامناسب هستند.

مثال ۱.۴. با فرض این که، جهت صفر در شرق و خلاف جهت عقربه‌های ساعت به عنوان جهت چرخش مثبت اختیار شود، فرض کنید دو پرند به ترتیب در جهات 15° و 345° پرواز کنند. میانگین حسابی آنها 180° ، نقطه‌ای در غرب است در حالی که دو مشاهده اساساً نقاطی در شرق هستند (شکل ۴ را ملاحظه کنید).



شکل ۴. محاسبه میانگین حسابی نقاط، به اشتباه!

مثال ۲.۴. فرض کنید یک مجموعه داده جهت‌دار شامل چهار جهت

^۹Unimodal

واریانس نمونه s^2 ، که به میانگین نمونه بستگی دارد به جهت صفر داریم. و جهت چرخش نیز وابسته است. به این ترتیب، در برخورد با داده‌های جهت‌دار به جایگزینی برای معیارهای تمرکز و پراکندگی نیاز داریم.

یک اندازه‌گیری مناسب و معنی‌دار از میانگین جهت برای یک مجموعه از جهات که تک‌مُدی هستند، با تمرکز دادن جهات به سمت یک جهت واحد با استفاده از جهت بردار برآیند آن‌ها به دست می‌آید. محاسبه جهت بردار برآیند برای یک مجموعه داده جهت‌دار به شرح زیر انجام می‌شود:

فرض کنید $\alpha_1, \alpha_2, \dots, \alpha_n$ یک مجموعه از مشاهدات برحسب زاویه باشند. تبدیل مختصات قطبی به دکارتی را برای هر مشاهده در نظر بگیرید (رابطه (۲) را ملاحظه کنید). یعنی:

$$(\cos \alpha_i, \sin \alpha_i), \quad i = 1, \dots, n$$

بردار برآیند این n بردار واحد، با مجموع مؤلفه‌های آن‌ها به صورت زیر به دست می‌آید:

$$R_0 = (\sum \cos(\alpha_i), \sum \sin(\alpha_i)) = (C, S). \quad (۳)$$

فرض کنید

$$R = \|R_0\| = \sqrt{C^2 + S^2},$$

طول بردار برآیند R_0 را نشان دهد. همان‌طور که می‌دانیم، هر بردار را می‌توان به صورت عدد مختلط در نظر گرفت. به عبارت دیگر، برای بردار برآیند R_0 به صورت زیر داریم:

$$R_0 = \sum_{j=1}^n \cos(\alpha_j) + i \sum_{j=1}^n \sin(\alpha_j),$$

که در آن $i = \sqrt{-1}$ می‌باشد.

بنابراین، جهت بردار برآیند R_0 ، که به عنوان میانگین جهت دایره‌ای پیشنهاد شده است، $\bar{\alpha}_0$ ، به صورت زیر تعریف می‌شود:

$$\bar{\alpha}_0 = \arg\{R_0\},$$

که به آن، آرگومان یا شناسه R_0 می‌گویند.

بنابراین، با معادلات

$$\cos(\bar{\alpha}_0) = C/R, \quad \sin(\bar{\alpha}_0) = S/R, \quad (۴)$$

$$\bar{\alpha}_0 = \arctan^*(S/C).$$

با توجه به این که جهت بردار برآیند، $\bar{\alpha}_0$ ، در چه ناحیه‌ای (ربع-ویژه) قرار بگیرد، داریم:

$$\bar{\alpha}_0 = \arctan^*(S/C) = \begin{cases} \arctan(S/C) & \text{if } C > 0, S \geq 0, \\ \arctan(S/C) + \pi & \text{if } C < 0, \\ \arctan(S/C) + 2\pi & \text{if } C \geq 0, S < 0, \\ Undefined & \text{if } C = 0, S = 0. \end{cases} \quad (۵)$$

به دنبال تعریف نسبتاً پیچیده «ربع-ویژه» معکوس تانژانت، این حقیقت که $\tan(\theta) = \tan(\theta + \pi)$ ضروری است، به طوری که دو معکوس برای هر θ داده شده وجود دارد. می‌دانیم که $\arctan$ مقادیری را در $(-\pi/2, \pi/2)$ می‌گیرد در حالی که تعریف (۵) معکوس منحصر به فردی را به سبب علامت‌های C و S روی $[0, 2\pi)$ فراهم می‌کند. گزاره زیر نشان می‌دهد که $\bar{\alpha}_0$ نشان دهنده مرکز برای مجموعه‌ای از داده‌ها است و به انتخاب مبدأ یا جهت چرخش بستگی ندارد.

گزاره ۳.۴. $\bar{\alpha}_0$ هموردای چرخشی^{۱*} است، یعنی اگر داده‌ها به یک مقدار معین انتقال داده شوند مقدار $\bar{\alpha}_0$ نیز به همان مقدار تغییر می‌کند.

اثبات. فرض کنید $\alpha_1, \alpha_2, \dots, \alpha_n$ دارای میانگین جهت $\bar{\alpha}_0$ باشند. می‌توان نشان داد که $(\alpha_1 + c, \alpha_2 + c, \dots, \alpha_n + c)$ دارای میانگین جهت $\bar{\alpha}_0 + c$ است. فرض کنید R'_0 بردار برآیند مجموعه جدیدی از مشاهدات، بعد از انتقال باشد. سپس داریم:

$$\begin{aligned} R'_0 &= \left(\sum_{i=1}^n \cos(\alpha_i + c), \sum_{i=1}^n \sin(\alpha_i + c) \right) \\ &= (C', S'), \end{aligned}$$

سپس

$$\begin{aligned} C' &= \sum_{i=1}^n \cos(\alpha_i + c) \\ &= \sum_{i=1}^n (\cos(\alpha_i) \cdot \cos(c) - \sin(\alpha_i) \cdot \sin(c)). \end{aligned}$$

^{۱*} Rotationally equivariant

۵ فاصله دایره‌ای و معیار پراکندگی

از رابطه (۳) داریم:

$$= C \cos(c) - S \sin(c),$$

از رابطه (۴) داریم:

$$= R \cos(\bar{\alpha}_0) \cos(c) - R \sin(\bar{\alpha}_0) \sin(c)$$

$$= R \cos(\bar{\alpha}_0 + c).$$

به طور مشابه:

$$S' = R \sin(\bar{\alpha}_0 + c),$$

$$R' = \|R'_0\| = \sqrt{C'^2 + S'^2}$$

$$= R \sqrt{\cos(\bar{\alpha}_0 + c)^2 + \sin(\bar{\alpha}_0 + c)^2}$$

$$= R = \sqrt{C^2 + S^2},$$

از این رو

$$C'/R' = \cos(\bar{\alpha}_0 + c), \quad S'/R' = \sin(\bar{\alpha}_0 + c)$$

$$\Rightarrow \tan(\bar{\alpha}_0 + c) = \frac{S'}{C'}$$

$$\Rightarrow \bar{\alpha}_0 + c = \arctan\left(\frac{S'}{C'}\right),$$

و در نتیجه اثبات می‌شود. □

به همین ترتیب، می‌توان بررسی کرد که $\bar{\alpha}_0$ نسبت به تغییرات در جهت چرخش نیز هموردا است. یعنی زمانی که حرکت در جهت عقربه‌های ساعت را به خلاف جهت عقربه‌های ساعت تبدیل می‌کنیم به طوری که α به $(2\pi - \alpha)$ تغییر می‌یابد، $\bar{\alpha}_0$ نیز به $(2\pi - \bar{\alpha}_0)$ تبدیل می‌شود.

تعریف ۴.۴. اگر بردار برآیند دارای طول مثبت باشد، جهت آن، $\bar{\alpha}_0$ به عنوان میانگین جهت دایره‌ای در نظر گرفته می‌شود و در صورتی که بردار برآیند به طول صفر باشد، میانگین جهت وجود ندارد.

زمانی که طول بردار برآیند صفر است، مربوط به موقعیتی است که داده‌ها به طور مساوی یا یکنواخت روی دایره پخش می‌شوند. در این مورد، واضح است که داده‌ها میانگین جهت را نشان نمی‌دهند.

نه تنها جهت بردار برآیند R ، $\bar{\alpha}_0$ ، میانگین جهت را فراهم می‌کند، بلکه طول آن

$$R = \|R_0\|,$$

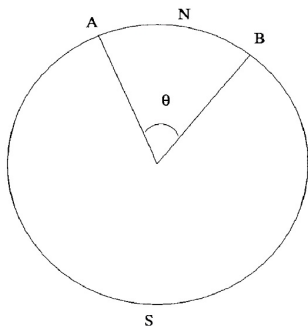
از این جهت که داده‌ها چگونه حول این میانگین متمرکز شده‌اند نیز، می‌تواند یک اندازه مفید برای داده‌های تک‌مدی باشد. از این رو، اگر تمام زوایا (بردارهای واحد) در یک نقطه متمرکز شده باشند R می‌تواند به بزرگی n باشد. به عکس، اگر داده‌ها به طور مساوی روی دایره پخش شده باشند تمرکز را نشان نمی‌دهند لذا R می‌تواند به کوچکی صفر باشد. حال با یک اندازه‌گیری مناسب، «فاصله» را روی دایره نشان می‌دهیم.

یک اندازه‌گیری معقول از «فاصله دایره‌ای» بین هر دو نقطه گرفتن کوچکترین کمان بین نقاط روی محیط دایره است. به عنوان مثال، برای هر دو زاویه α و β :

$$d_0(\alpha, \beta) = \min(\alpha - \beta, 2\pi - (\alpha - \beta))$$

$$= \pi - |\pi - |\alpha - \beta||, \quad (6)$$

یعنی، فاصله بین A و B می‌تواند طول کمان $\widehat{ANB}$ یا طول کمان $\widehat{ASB}$ باشد (شکل ۶ را ملاحظه کنید). اما چون کمان $\widehat{ANB}$ کوچکتر از کمان $\widehat{ASB}$ است، فاصله دایره‌ای را طول کمان $\widehat{ANB}$ تعریف می‌کنیم. واضح است که فاصله هیچ دو نقطه‌ای روی محیط دایره نمی‌تواند بیشتر از π یا فاصله دایره‌ای در تعریف (۶) باشد. به عبارت دیگر، فاصله دایره‌ای همیشه در بازه $[0, \pi]$ قرار می‌گیرد.



شکل ۶. فاصله دایره‌ای d_0 ، طول کمان $\widehat{ANB}$ است.

تعریف دیگری از «فاصله دایره‌ای» بین نقاط A و B در شکل ۶ به صورت زیر است:

$$d(\alpha, \beta) = 1 - \cos(\alpha - \beta), \quad (7)$$

که α و β به ترتیب نشان دهنده زوایای مربوط به A و B هستند. اثبات. فرض کنید

$$u_i = (\cos(\alpha_i), \sin(\alpha_i)) = (x_i, y_i),$$

اگر θ زاویه بین نقاط A و B باشد، واضح است که $d(\alpha, \beta)$ یک تابع صعودی یکنواخت از θ است و زمانی که $\theta = 0$ باشد، $d(\alpha, \beta) = 0$ و

$$v = (a, b),$$

و زمانی که $\theta = \pi$ باشد، مقدار $d(\alpha, \beta)$ به ۲ افزایش می یابد.

که در آن $a^2 + b^2 = 1$. با فرض بر این که، α_i و β به ترتیب زوایای بردارهای u_i و v باشند و زاویه بین این بردارها را با $\theta_i = \beta - \alpha_i$ نشان دهیم، داریم:

$$\begin{aligned} d(v, u_i) &= 1 - \cos(\theta_i) = 1 - \cos(\beta - \alpha_i) \\ &= 1 - (ax_i + by_i) = 1 - (v \cdot u_i), \end{aligned}$$

که در آن، حاصل ضرب داخلی از بردارهای v و u_i است. حال، با استفاده از تعریف (۸) داریم:

$$D_v(u_1, u_2, \dots, u_n) = n - \sum_{i=1}^n \cos(\theta_i) = n - \sum_{i=1}^n (ax_i + by_i).$$

ار آنجایی که $a^2 + b^2 = 1$ ، صفر قرار دادن مشتق D_v نسبت به یکی از مؤلفه های a و b کافی است. لذا مشتق D_v نسبت به a را محاسبه و برابر با صفر قرار می دهیم:

$$0 = \frac{dD_v}{da} = - \sum_{i=1}^n \left\{ x_i + \frac{db}{da} y_i \right\},$$

با توجه به انتخاب v ، چون $a^2 + b^2 = 1$ داریم:

$$2a(da) + 2b(db) = 0,$$

$$\frac{db}{da} = -\frac{a}{b}, \quad (9)$$

لذا با استفاده از رابطه (۹) داریم:

$$0 = \frac{dD_v}{da} = - \sum_{i=1}^n \left\{ x_i - \frac{a}{b} y_i \right\}.$$

از این رو:

$$\frac{a}{b} = \frac{\sum_{i=1}^n x_i}{\sum_{i=1}^n y_i} = \frac{C}{S}$$

چون $a^2 + b^2 = 1$ داریم:

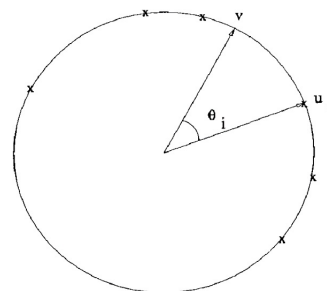
$$\frac{C^2}{S^2} = \frac{a^2}{1 - a^2}$$

که دلالت بر:

$$a = \frac{C}{R}, \quad b = \frac{S}{R},$$

مجموعه ای از بردارهای واحد را به صورت $\{u_i, i = 1, \dots, n\}$ در نظر بگیرید و فرض کنید که $D_v(u_1, u_2, \dots, u_n)$ پراکندگی نمونه را با توجه به بردار واحد دلخواه $v = (a, b)$ نشان دهد. اگر θ زاویه بین u_i و بردار دلخواه v باشد (شکل ۷ را ملاحظه کنید)، $0 \leq \theta_i \leq \pi$ ، فاصله دایره ای $d(v, u_i) = 1 - \cos(\theta_i)$ را با استفاده از تعریف (۷) داریم. بنابراین:

$$D_v(u_1, u_2, \dots, u_n) = \sum_{i=1}^n d(v, u_i) = n - \sum_{i=1}^n \cos(\theta_i), \quad (8)$$



شکل ۷. «فاصله دایره ای» بین V و مشاهدات u_i روی دایره

مشابه موارد خطی (که پراکندگی اطراف یک مقدار مانند c نامیده می شود، یعنی $\sum (x_i - c)^2$. این مقدار زمانی که $c = \bar{x}$ و کمینه می شود و به عنوان تعریفی برای واریانس به کار می رود)، می توان یک نقطه مانند v روی محیط دایره یافت، به طوری که $D_v(u_1, u_2, \dots, u_n)$ را کمینه کند. گزاره ۱.۵، $(n - R)$ را به عنوان یک اندازه از پراکندگی مشابه واریانس نمونه s^2 برای داده های خطی، تعریف می کند.

گزاره ۱.۵. $D_v(u_1, u_2, \dots, u_n)$ زمانی که v بردار برآیند نرمالیده^{۱۱}

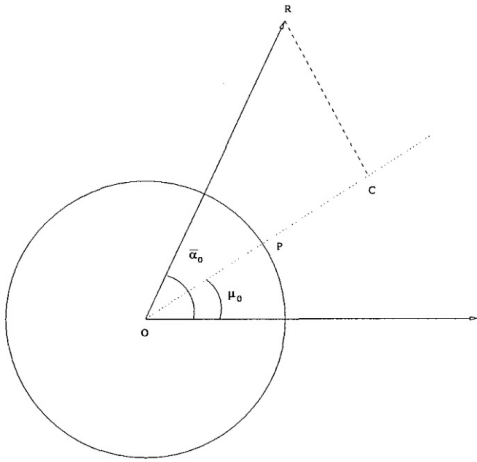
به صورت زیر است، کمینه می شود

$$v^* = \left(\frac{C}{R}, \frac{S}{R} \right).$$

لذا اندازه مربوط به پراکندگی نمونه به صورت زیر محاسبه می شود:

$$D_{v^*} = n - R.$$

^{۱۱} Normalized



شکل ۸. تصویر بردار برآیند روی بردار قطبی

قضیه ۲.۵. اگر $\bar{\alpha}_0$ جهت بردار برآیند برای نمونه $\alpha_1, \alpha_2, \dots, \alpha_n$ باشد، داریم:

$$\sum_{i=1}^n \sin(\alpha_i - \bar{\alpha}_0) = 0$$

$$\sum_{i=1}^n \cos(\alpha_i - \bar{\alpha}_0) = R.$$

اثبات. به این شرح است که:

$$\begin{aligned} \sum_{i=1}^n \sin(\alpha_i - \bar{\alpha}_0) &= \sum_{i=1}^n \{\sin(\alpha_i) \cos(\bar{\alpha}_0) - \cos(\alpha_i) \sin(\bar{\alpha}_0)\} \\ &= Y \cos(\bar{\alpha}_0) - X \sin(\bar{\alpha}_0) \\ &= R \cos(\bar{\alpha}_0) \sin(\bar{\alpha}_0) - R \cos(\bar{\alpha}_0) \sin(\bar{\alpha}_0) \\ &= 0 \end{aligned}$$

و برای $\sum_{i=1}^n \cos(\alpha_i - \bar{\alpha}_0) = R$ به طور مشابه عمل می‌شود. □

تذکر ۳.۵. این نتایج را می‌توان با برخی شباهت‌ها با داده‌های خطی ارائه کرد. به طور مثال، مجموعه $\{x_1, x_2, \dots, x_n\}$ با میانگین نمونه $\bar{x}$ و میانگین جامعه μ را در نظر بگیرید. توجه داشته باشید که $D_{v^*} = (n - R)$ متناظر با $\sum (x_i - \bar{x})^2$ ، در حالی که $D_P = (n - V_0)$ متناظر با $\sum (x_i - \mu)^2$ پراکندگی اطراف میانگین حقیقی است. قضیه ۲.۵ این حقیقت را یادآوری می‌کند که $\sum (x_i - \bar{x}) = 0$ و $\sum (x_i - \mu) = 0$ واریانس نمونه را نشان می‌دهد.

دارد. از این‌رو، D_v زمانی مینیمم می‌شود که $v = v^* = (\frac{C}{R}, \frac{S}{R})$ بردار برآیند نرمالیده باشد. پراکندگی نمونه دایره‌ای با توجه به این بردار به صورت زیر است:

$$\begin{aligned} D_{v^*} &= \sum_{i=1}^n d(v^*, u_i) = n - \sum_{i=1}^n \left\{ \frac{C}{R} x_i + \frac{S}{R} y_i \right\} \\ &= n - \left(\frac{C}{R} C + \frac{S}{R} S \right) \\ &= n - R. \end{aligned}$$

□

اگر مقدار R نزدیک به صفر باشد به این معنی است که اندازه پراکندگی بزرگ است و در صورتی که مقدار R نزدیک به n باشد، بیان می‌کند که مجموعه‌ای از مشاهدات دارای پراکندگی کوچک یا تمرکز بیشتر به سمت مرکز هستند. لازم به ذکر است که $(n - R)$ پراکندگی مربوط به نمونه را اندازه می‌گیرد.

و اگر میانگین جهت جامعه، μ_0 ، را که جهت قطبی نیز نامیده می‌شود به دست آوریم سپس «برداری قطبی» را با $P = (\cos(\mu_0), \sin(\mu_0))$ نشان دهیم، اندازه پراکندگی اطراف این میانگین جهت معلوم را می‌توان به صورت زیر ارائه کرد به طوری که:

$$D_P = n - \sum_{i=1}^n \cos(\alpha_i - \mu_0) = n - V_0$$

که

$$\begin{aligned} V_0 &= \sum_{i=1}^n \{\cos(\alpha_i) \cos(\mu_0) + \sin(\alpha_i) \sin(\mu_0)\} \\ &= X \cos(\mu_0) + Y \sin(\mu_0) \\ &= R \cos(\bar{\alpha}_0) \cos(\mu_0) + R \sin(\bar{\alpha}_0) \sin(\mu_0) \\ &= R \cos(\bar{\alpha}_0 - \mu_0) \\ &= R.C \quad (C = \cos(\bar{\alpha}_0 - \mu_0)), \end{aligned}$$

V_0 متناظر با طول OC در شکل ۸ است، که طول تصویر بردار برآیند نمونه متمایل به جهت قطبی P را نشان می‌دهد. چون $C = \cos(\bar{\alpha}_0 - \mu_0) \leq 1$ ، لذا $V_0 \leq R$ و برابر است اگر و فقط اگر $\bar{\alpha}_0 = \mu_0$ باشد. این نتیجه نیز در نیمه دوم قضیه زیر تکرار شده است.

تذکر ۱.۶. گاهی اوقات داده‌ها در یک شکل گروهی رخ می‌دهند، که به آن‌ها داده‌های گروهی می‌گویند (به عنوان مثال داده‌های رودخانه *Kamthi* در جدول ۲ را ملاحظه کنید). در این مورد، فرض معمول بر این است که همه مشاهدات در یک فاصله و در نقطه میانی آن فاصله قرار دارند. اگر به عنوان مثال، n مشاهده اصلی به k کلاس با i امین کلاس دارای یک نقطه میانی α_i و فراوانی f_i ، $(\sum_{i=1}^k f_i = n)$ $i = 1, 2, \dots, k$ گروه‌بندی شده باشند، سپس داریم:

$$\bar{C}_n = \frac{1}{n} \sum_{i=1}^k f_i \cos(\alpha_i),$$

و

$$\bar{S}_n = \frac{1}{n} \sum_{i=1}^k f_i \sin(\alpha_i),$$

لذا

$$\bar{R} = \sqrt{\bar{C}_n^2 + \bar{S}_n^2}.$$

به طور مشابه گشتاورهای مثلثاتی مرتبه بالاتر به صورت زیر محاسبه می‌شوند:

$$\bar{C}_n(p) = \frac{1}{n} \sum_{i=1}^k f_i (\cos(p\alpha_i)),$$

و

$$\bar{S}_n(p) = \frac{1}{n} \sum_{i=1}^k f_i (\sin(p\alpha_i)).$$

برای بسیاری از اهداف کاربردی تصحیح گروه‌بندی نیاز نیست، مگر در مواردی که گروه‌بندی بسیار درشت است (با فاصله بیشتر از 45° در طول را می‌گویند).

تذکر ۲.۶. گاهی اوقات داده‌های جهت‌دار به جای این که به جهت اشاره کنند به «محور» اشاره می‌کنند، به عنوان مثال محور طولانی از ذرات در رسوبات یا محور نوری کریستال. چنین مشاهداتی از محورها به داده‌های محوری اشاره دارند. در این مورد، هر جهت θ با جهت مخالف آن $(\theta + \pi)$ تشخیص داده می‌شود. چنین داده‌هایی به شیوه «دو برابر زوایا» کنترل می‌شوند، یعنی تبدیل θ به 2θ که ابهام جهت‌دار بودن را برطرف می‌کند. «میانگین محوری» متناظر با داده‌های محوری $\theta_1, \theta_2, \dots, \theta_n$ به صورت زیر به دست می‌آید: میانگین جهت $\bar{\alpha}_0$ را از زوایای دو برابر شده $(\alpha_i = 2\theta_i, i = 1, 2, \dots, n)$ به دست آورده، سپس نصف آن را در نظر می‌گیریم (یعنی $\frac{1}{2}\bar{\alpha}_0$ یا به طور معادل $\frac{1}{2}\bar{\alpha}_0 + \pi$).

در صورتی که θ کوچک باشد، با استفاده از بخش دوم قضیه ۲.۵ در گام اول و بسط تیلور برای $\cos \theta$ در گام دوم، داریم:

$$\begin{aligned} \sin \theta &\approx \theta, \quad \cos \theta \approx 1 - \frac{\theta^2}{2} \\ 2(1 - \bar{R}) &= \frac{2}{n} \sum_{i=1}^n (1 - \cos(\alpha_i - \bar{\alpha}_0)) \\ &\cong \frac{1}{n} \sum_{i=1}^n (\alpha_i - \bar{\alpha}_0)^2 \\ &= S^2, \end{aligned}$$

که این منجر به تعریف $2(1 - \bar{R})$ به عنوان یک «فاصله دایره‌ای» می‌شود.

۶ گشتاورهای بالاتر

همان‌طور که قبلاً گفته شد، C و S را به صورت مجموع مؤلفه‌های $\cos$ و $\sin$ تعریف کردیم. حال می‌توان میانگین آن‌ها را به صورت زیر در نظر گرفت:

$$\bar{C}_n = \frac{1}{n} \sum_{j=1}^n \cos(\alpha_j), \quad \bar{S}_n = \frac{1}{n} \sum_{j=1}^n \sin(\alpha_j),$$

و چون $e^{i\alpha} = \cos \alpha + i \sin \alpha$ ، لذا داریم:

$$\bar{C}_n + i\bar{S}_n = \frac{1}{n} \sum_{j=1}^n e^{i\alpha_j}.$$

به طور مشابه می‌توانیم گشتاورهای نمونه مرتبه بالاتر را با گرفتن توان بالاتر از $\{e^{i\alpha_j}\}$ و میانگین آن‌ها تعریف کنیم:

$$e^{ip\alpha} = \cos(p\alpha) + i \sin(p\alpha),$$

داریم:

$$\begin{aligned} \frac{1}{n} \sum_{j=1}^n e^{(i\alpha_j)^p} &= \frac{1}{n} \sum_{j=1}^n e^{ip\alpha_j} \\ &= \frac{1}{n} \sum_{j=1}^n \cos(p\alpha_j) + i \frac{1}{n} \sum_{j=1}^n \sin(p\alpha_j) \\ &= \bar{C}_n(p) + i\bar{S}_n(p), \end{aligned} \quad (10)$$

که $(\bar{C}_n(p), \bar{S}_n(p))$ گشتاورهای مثلثاتی مرتبه p ام براساس نمونه نامیده می‌شوند.

بنابراین:

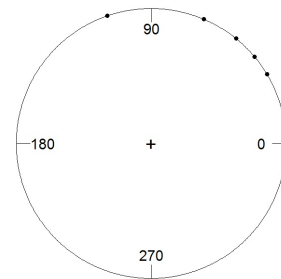
$$\bar{S}(1) = 0.63, \quad \bar{C}(1) = 0.55$$

جدول ۲: توزیع فراوانی مشخصات جغرافیایی لایه‌های متقاطع در سه سطح از رودخانه *Kamthi* (سنگپتا و راثو[۸]).

بالای <i>Kamthi</i>	وسط <i>Kamthi</i>	پایین <i>Kamthi</i>	مشخصات جغرافیایی (برحسب درجه)
۷۵	۵۰	۱۴	۰-۱۹
۷۵	۶۲	۱۴	۲۰-۳۹
۱۵	۳۳	۱۱	۴۰-۵۹
۲۵	۹	۱۳	۶۰-۷۹
۷	۱	۹	۸۰-۹۹
۳	۳	۱۶	۱۰۰-۱۱۹
۳	۰	۰	۱۲۰-۱۳۹
۰	۰	۴	۱۴۰-۱۵۹
۰	۰	۰	۱۶۰-۱۷۹
۰	۰	۳	۱۸۰-۱۹۹
۲۱	۲	۴	۲۰۰-۲۱۹
۸	۸	۰	۲۲۰-۲۳۹
۲۴	۰	۰	۲۴۰-۲۵۹
۱۶	۱۱	۰	۲۶۰-۲۷۹
۳۶	۵	۶	۲۸۰-۲۹۹
۷۵	۲۰	۷	۳۰۰-۳۱۹
۹۰	۵۳	۱	۳۲۰-۳۳۹
۱۰۷	۴۱	۲۱	۳۴۰-۳۵۹

همچنین، نمودار پراکنش و نمودار میانگین جهت برای داده‌های دایره‌ای را می‌توان به طور جداگانه رسم کرد. برای این منظور از دستور *circ.plot* در بسته *CircStats* استفاده می‌کنیم:

```
> circ.plot(rad(x))
```



شکل ۹. نمودار پراکنش داده‌های دایره‌ای در نرم‌افزار R

۷ نرم افزار R

مثال زیر نشان می‌دهد که چگونه می‌توان از بسته *CircStats* برای محاسبه میانگین نمونه، پراکندگی و گشتاورهای مثلثاتی استفاده کرد.

مثال ۱.۷. فرض کنید داده‌ها به صورت زیر جمع آوری شده باشند:

$$51^\circ, 67^\circ, 40^\circ, 109^\circ, 31^\circ, 358^\circ$$

و بخواهیم میانگین جهت، اندازه پراکندگی و گشتاور مثلثاتی نمونه مرتبه اول را محاسبه کنیم.

برای این منظور به ترتیب از برنامه‌های *circ.mean*، *circ.disp* و *trig.moment* در بسته *CircStats* استفاده می‌کنیم:

```
> x =c(51,67,40,109,31,358)
```

```
> circ.mean(rad(x))
```

```
[1] 0.8487044
```

```
> circ.disp(rad(x))
```

```
      n      r      rbar      var
```

```
1 6 5.018619 0.8364365 0.1635635
```

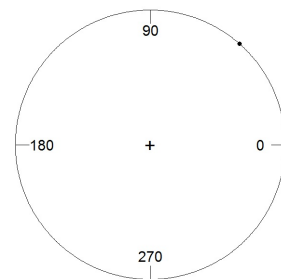
```
> trig.moment(rad(x),1)
```

```
      mu.p      rho.p      cos.p      sin.p
```

```
0.8487044 0.8364365 0.5528477 0.6276826
```

در شکل ۹، داده‌ها به صورت نقاطی روی دایره واحد بر حسب درجه نشان داده شده‌اند و در شکل ۱۰، میانگین جهت داده‌ها نیز بر حسب درجه روی دایره واحد نشان داده شده است.

```
> circ.plot(circ.mean(rad(x)))
```



شکل ۱۰. نمایش دایره‌ای میانگین جهت داده‌ها در نرم‌افزار R

۸ نرم افزار Matlab

در این بخش، مثال پیشین با استفاده از جعبه ابزار *circstats* در نرم‌افزار Matlab نیز مورد بررسی قرار گرفته و میانگین جهت، پراکندگی و گشتاور مثلثاتی محاسبه شده است. این نتایج را می‌توان با هم مقایسه کرد.

```
>> circ_ang2rad(data)
```

```
ans = 0.8901
```

همان‌طور که ملاحظه می‌کنید، مقادیر محاسبه شده با استفاده از نرم‌افزار R و Matlab تفاوتی نمی‌کنند تنها تفاوت در شیوه نمایش است. همچنین در نرم‌افزار Matlab اندازه بردار برآیند، R ، به طور مستقیم با استفاده از دستور *circ_r* به دست می‌آید، در حالی که در نرم افزار R علاوه بر این که به وسیله دستور *est.rho* به طور جداگانه محاسبه می‌شود، با استفاده از دستور *circ.summary* همراه با میانگین نیز می‌توان به دست آورد.

۹ نتیجه‌گیری

مجموعه‌ای از مشاهدات را که بر اساس زاویه جمع‌آوری می‌شوند، داده‌های جهت‌دار می‌نامند. این زوایا یا جهات را در فضای دو بُعدی (دایره) یا سه بُعدی (کره) می‌توان نشان داد، که در اینجا تنها جهات دو بُعدی را مورد بررسی قرار دادیم. جهات‌های دو بُعدی را می‌توان به صورت زوایای اندازه‌گیری شده با توجه به انتخاب مناسب «جهت صفر» و «جهت چرخش» اختیار کرد.

از آنجایی که یک جهت دارای مقدار و اندازه نیست، می‌توان به صورت نقاطی روی محیط یک دایره واحد ($r = 1$) متمرکز در مبدأ یا به صورت بردارهای واحد نشان داد. به دلیل این نمایش دایره‌ای، این نوع از مشاهدات را داده‌های دایره‌ای نیز می‌نامند.

به طور خلاصه، برای مجموعه داده شده از زوایا، ابتدا مختصات قطبی زوایا را به مختصات دکارتی تبدیل و سپس بردار برآیند آنها را همان‌طور که ارائه کردیم، محاسبه می‌کنیم. حال، جهت بردار برآیند یعنی $\bar{\alpha}_0$ را با استفاده از تعریف (۵) به دست می‌آوریم، که نشان‌دهنده میانگین جهت دایره‌ای برای مجموعه داده شده از زوایا است و می‌توان با استفاده از برنامه *circ.mean* در بسته *CircStats* محاسبه کرد.

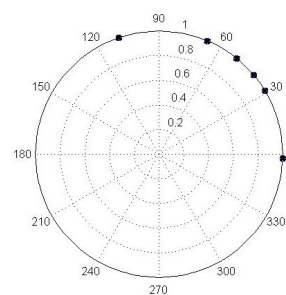
واضح است که وقتی $R = 0$ است، یعنی زمانی که $C = S = 0$ ، میانگین جهت را نمی‌توان تعریف کرد. این مربوط به موقعیتی است که داده‌ها به طور مساوی یا یکنواخت روی دایره پخش می‌شوند. با توجه به تعریف بردار برآیند و فاصله دایره‌ای، پراکندگی نمونه را به صورت زیر می‌توان تعریف کرد:

$$s^2 = n - R.$$

```
1.1694
0.6981
1.9024
0.5411
6.2483
>> alpha_bar=circ_mean(ans)
alpha_bar = 0.8487
>> R=circ_r(ans)
R = 0.8364
>> S=circ_var(ans)
S = 0.1636
>> s=circ_std(ans)
s = 0.5720
>> mp=circ_moment(ans,[],1)
mp = 0.5528 + 0.6277i
```

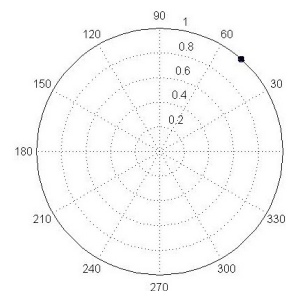
با استفاده از دستور *circ_plot* در جعبه ابزار *circstats* در نرم‌افزار Matlab، می‌توان داده‌ها و میانگین جهت داده‌ها را همانند مثال قبل روی دایره واحد نشان داد.

```
>> circ_plot(data)
```



شکل ۱۱. نمایش دایره‌ای داده‌ها در نرم‌افزار Matlab

```
>> circ_plot(alpha_bar)
```



شکل ۱۲. نمایش دایره‌ای میانگین جهت داده‌ها در نرم‌افزار

Matlab

مراجع

- [1] Berens, Ph. (2009). " *CircStat: A Matlab Toolbox for Circular Statistics*", Journal of Statistical Software, Vol. 31.
- [2] Fisher, N. I. (1993). *Statistical Analysis of Circular Data*. Cambridge: Cambridge University Press.
- [3] Jammalamadaka, S. R., Bhadra, N., Chaturvedi, D., Kuttty, T. K., Majumdar, P. P., And Poduval, G. (1986), Functional assessment of knee and ankle during level walking, In Krishnan, T., editor, *Data Analysis in Life Science*, pages 21-54. Indian Statistical Institute, Calcutta, India.
- [4] Jammalamadaka, S. R. and Sarma, Y. R. (1988), A correlation coefficient for angular variables, In Matusita, K., editor, *Statistical Theory and Data Analysis II*, pages 349-364. North Holland, Amsterdam.
- [5] Jammalamadaka, S. Rao. & SenGupta, A. (2001). *Topics in circular statistics*, Vol. 5, pages 1–24.
- [6] Mardia, K. V. (1972), *Statistics of Directional Data*, Academic Press, New York.
- [7] Schmidt-Koenig, K. (1963), On the role of loft, the distance and site of release in pigeon homing (the "cross-loft experiment"), *Biol. Bull.*, 125, 154-164.
- [8] Sengupta, S. and Rao, J. S. (1967), Statistical analysis of crossbedding azimuths from the Kamthi formation around Bheemaram, Pranhita Godavari Valley, *Sankhya*, 28, 165-174.