پروژه اول درس برنامه‌سازی کامپیوتر

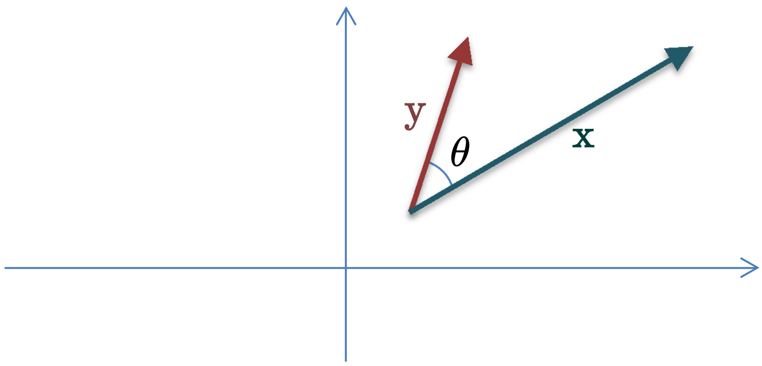
در جبر خطی، ضرب داخلی[1] دو بردار در فضای دو بعدی اقلیدسی عدد خاصی را به آن دو بردار نسبت می‌دهد که کاربردهای گوناگونی در آنالیز برداری در حوزه‌های مختلف فیزیک و مهندسی دارد.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad , \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$\mathbf{x} \bullet \mathbf{y} = x_1 y_1 + x_2 y_2 = \|\mathbf{x}\| \|\mathbf{y}\| \cos\theta$$

$$\|\mathbf{x}\| = \sqrt{{x_1}^2 + {x_2}^2} \quad , \quad \|\mathbf{y}\| = \sqrt{{y_1}^2 + {y_2}^2}$$

که در آن θ زاویه‌ی بین دو بردار و ||x|| و ||y|| به ترتیب نُرم اقلیدسی[2] (اندازه) دو بردار هستند. نمایش هندسی بردارها در فضای دوبعدی قابل تصور است و می‌توان آن‌ها را به سادگی در صفحه نمایش داد:



در فضای سه‌بعدی نیز روابطی مشابه برقرارند، تنها با این تفاوت که بعد سومی به آن‌ها اضافه شده است:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad , \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

$$\mathbf{x} \bullet \mathbf{y} = x_1 y_1 + x_2 y_2 + x_3 y_3 = \|\mathbf{x}\| \|\mathbf{y}\| \cos\theta$$

$$\|\mathbf{x}\| = \sqrt{{x_1}^2 + {x_2}^2 + {x_3}^2} \quad , \quad \|\mathbf{y}\| = \sqrt{{y_1}^2 + {y_2}^2 + {y_3}^2}$$

---

[1] Inner product
[2] Euclidian Norm

شباهت دو بردار : فاصله[3] و شباهت زاویه‌ای[4]:

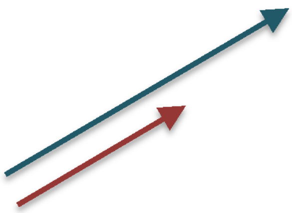
در جبر خطی از معیارهای مختلفی برای به دست آوردن شباهت دو بردار استفاده می‌شود. یکی از این معیارها استفاده از مفهوم **فاصله** است که اینگونه تعریف می‌شود:

$$d(\mathbf{x},\mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$$

طبیعتا وقتی این فاصله برابر صفر خواهد شد که دو بردار هم‌اندازه و هم‌جهت باشند، یعنی با هم کاملا برابر باشند. اما معیار دیگری نیز برای بررسی میزان شباهت دو بردار وجود دارد که به **شباهت زاویه‌ای** معروف است. طبیعتا هر چه زاویه دو بردار کمتر باشد، دو بردار به هم شبیه‌تر هستند. حالا به خوبی می‌دانیم که این زاویه را می‌توان از رابطه‌ی زیر به دست آورد:

$$\theta = \cos^{-1}\left(\frac{\mathbf{x} \bullet \mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}\right) = \cos^{-1}\left(\frac{x_1y_1 + x_2y_2 + x_3y_3}{\sqrt{{x_1}^2 + {x_2}^2 + {x_3}^2}\ \sqrt{{y_1}^2 + {y_2}^2 + {y_3}^2}}\right)$$

وقتی زاویه‌ی دو بردار صفر باشد، به طور کامل مشابه در نظر گرفته می‌شوند. بسیار مهم است به تفاوت این مفهوم با مفهوم فاصله دقت داشته باشید: در شباهت زاویه‌ای فقط جهت مهم است و اندازه (یا همان نرم) ملاک نیست. بنابراین بردارهای زیر با معیار شباهت زاویه‌ای کاملا مشابه تلقی می‌شوند، اما به معیار فاصله خیر. هر کجا که اندازه برای ما اهمیت نداشته باشد می‌توانیم از این معیار استفاده کنیم.



بنابراین می‌توان معیار شباهت زاویه‌ای را اینگونه تعریف کرد:

$$\text{Angular Similariy } = \cos\theta$$

- اگر این عدد ۱ باشد دو بردار کاملا مشابه (یعنی هم‌جهت) قلمداد می‌شوند.
- اگر صفر باشد می‌گوییم بر هم عمود هستند. با این مفهوم به خوبی آشنا هستیم، یعنی زاویه بین دو بردار ۹۰ درجه بوده و طبیعتا ضرب داخلی دو بردار برابر صفر است.
- اگر برابر ۱- باشد کاملا غیر مشابه در نظر گرفته می‌شوند.

طبیعی است که هر چه این معیار به ۱ نزدیک‌تر باشند، دو بردار شبیه‌تر هستند.

---

[3] Distance
[4] Angular Similarity

فضاهای N بعدی[۵]:

ما درک روشنی از فضاهای دوبعدی و سه‌بعدی داریم. با نگرش فیزیک کلاسیک، جهان اطراف ما یک فضای سه‌بعدیِ مکانی است. درک این فضا برای ذهن ما بسیار ساده است؛ اما وقتی ابعاد فضا بیشتر شود، دیگر تصوری هندسی و تصویری وجود ندارد. با این حال از نظر ریاضی روابط گفته‌شده با همان تعاریف همچنان برقرار و معنی‌دار هستند. در واقع می‌توان این روابط را در حالت کلی به فضاهای N بعدی تعمیم داد:

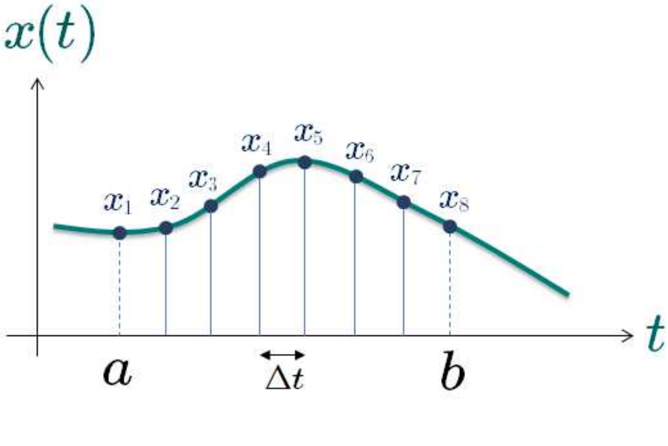
$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_N \end{bmatrix} \quad , \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_N \end{bmatrix}$$

$$\mathbf{x} \bullet \mathbf{y} = \sum_{k=1}^{N} x_k y_k = \|\mathbf{x}\| \|\mathbf{y}\| \cos\theta$$

$$\|\mathbf{x}\| = \sqrt{{x_1}^2 + {x_2}^2 + {x_3}^2 + \cdots + {x_N}^2} \quad , \quad \|\mathbf{y}\| = \sqrt{{y_1}^2 + {y_2}^2 + {y_3}^2 + \cdots + {y_N}^2}$$

به زودی متوجه خواهید شد که این تعمیم چه کاربردهای متنوعی می‌تواند داشته باشد.

هر سیگنال (تابع) پیوسته را می‌توان با مقادیری گسسته از نمونه‌های آن نمایش داد. به این کار نمونه‌برداری[۶] می‌گویند. فرض کنید $x(t)$ سیگنال ولتاژ حاصل از یک میکروفون باشد. این سینگال در واقع تابعی بر حسب زمان است. شکل زیر نشان می‌دهد که چگونه ۸ نمونه از این سیگنال در بازه زمانی a تا b برداشته شده و در یک بردار $\mathbf{x}$ قرار داده شده است.



$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \end{bmatrix}$$

---

[5] N-Dimensional Space

[6] Sampling

بردار x یک بردار در فضای 8 بعدی قلمداد می‌شود که در واقع نمونه‌هایی از تابع پیوسته است. اگر تعداد نمونه‌ها ۱۰۰۰ عدد باشد، می‌گوییم که بردار به یک فضای ۱۰۰۰ بعدی تعلق دارد. شما قبلا در طی درس با این مفهوم آشنا شده‌اید، آنجا که مقادیری گسسته از یک تابع خاص را با استفاده از کتابخانه‌ی `numpy` محاسبه می‌کردیم و با استفاده از کتابخانه‌ی `matplotlib` آن را رسم می‌کردیم.

به Δt دوره‌ی تناوب نمونه‌برداری و به عکس آن فرکانس نمونه‌برداری گفته می‌شود:

$$\text{Sampling Period}: \quad \Delta t$$
$$\text{Sampling Frequency}: \quad f_s = \frac{1}{\Delta t}$$

طبیعی است که برای اینکه بردار گسسته (نمونه‌برداری‌شده) نمایش صحیحی از سیگنال اصلی پیوسته را ارائه کند، Δt باید به اندازه کافی کوچک باشد، یعنی نمونه‌ها به اندازه کافی به هم نزدیک باشند. طی ترم‌های آینده در درس تجزیه و تحلیل سیستم‌ها خواهید دید که این مقدار حداقل چقدر باید باشد و چگونه به دست می‌آید.

حالا با تمام این توضیحاتی که ارائه شد، می‌توان شباهت زاویه‌ای دو تابع نمونه‌برداری‌شده را به سادگی به دست آورد. توجه داشته باشید که با این تعریف، بردارهای حاصل از نمونه‌برداری دو تابع $\sin(2t)$ و $5\sin(2t)$ کاملا مشابه هستند. واضح است که این دو تابع شکل کاملا یکسانی دارند و فقط اندازه شان با هم فرق دارد، در واقع یکی ضریبی از دیگری است. (در جبر خطی به چنین بردارهایی وابسته‌ی خطی[۷] گفته می‌شود، که نقطه‌ی مقابل مستقل خطی[۸] است.)

[7] Linear dependent
[8] Linear independent

شرح پروژه:

۱- با استفاده از کتابخانه‌ی **numpy** برنامه‌ای بنویسید که شباهت زاویه‌ای دو تابع زیر را در هر یک از بازه‌های داده‌شده با استفاده از روابطی که گفته شد به دست آورید (از بردارهایی با بعد ۱۰۰۰ استفاده کنید). حتما تابعی بنویسید که شباهت را محاسبه کند و برگرداند. سپس تابع را در هر مورد فراخوانی کرده و میزان شباهت را به دست آورید. با استفاده از **matplotlib** در هر مورد نمودارهای مربوطه را هم رسم کنید. نهایتا، حاصل از معیار شباهت را با چیزی که از نگاه کردن به نمودار برداشت می‌کنید توضیح دهید.

$$f(t) = 1.7t^2 \quad , \quad g(t) = e^t - 1$$

در بازه‌های [0,1] و [0,10]

۲- برنامه‌ای بنویسید که مقدار a را برای اینکه دو تابع زیر بیشترین شباهت را به هم داشته باشند، پیدا کند.

$$f(t) = \frac{1}{\sqrt{1+t^2}} \quad , \quad g(t) = e^{-at^2}$$

در بازه‌ی [0,5]

۳- چهار فایل صوتی با فرمت wav. ضمیمه شده‌اند در در هریک از آنها موتیف ابتدایی از سمفونی شماره‌ی پنج اثر لودویگ وان بتهوون (سازنده چندی از بزرگ‌ترین آثار موسیقی) هر کدام با یک ساز مختلف (ویلن، ویولا، پیانو و هورن فرانسوی) اجرا شده است. این چهار فایل را در یک برنامه پایتون باز کنید و بردار عددی هر یک را در یک متغیر (آرایه) قرار دهید.

الف) طول بردارها چقدر است؟ فرکانس نمونه‌برداری از این صداها چقدر است؟ (عدد آشنایی که حتما روی فایل‌های صوتی دیده‌اید). دوره تناوب نمونه‌برداری را به دست آورید.

ب) در برنامه‌ی خود نمودار سیگنال‌های صوتی را بر حسب زمان ترسیم کنید. با توجه به دوره‌ی تناوب نمونه‌برداری و طول بردار مربوط به سیگنال‌ها، باید بردار زمانی را برای رسم نمودار تشکیل دهید.

ج) در برنامه‌ی خود بردار مربوط به صدای ویلن را مبنا قرار داده و شباهت زاویه‌ای آن را با هریک از سه صدای دیگر محاسبه کنید. این نتایج چه می‌گوید؟ کدام یک از این صداها به صدای ویلن شبیه‌تر است؟ آیا این نتیجه با چیزی که می‌شنوید مطابقت دارد؟

راهنمایی - واردکردن فایل‌های صوتی wav. در پایتون:

```
import soundfile as SF
waveData,fs = SF.read('piano.wav')
```

که fs فرکانس نمونه‌برداری را می‌دهد. چون فایل‌ها به صورت استریو (دو کانال R/L) ذخیره شده‌اند، برای سادگی حتما کانال ۱ را استخراج کنید و با آن کار کنید:

```
x=waveData[:,1]
```