Aproximarea unei melodii folosind mai multe metode de calcul numeric

Popa Mircea Alex
Sîrghe Matei Ștefan
Ungureanu Robert

May 31, 2025

1 Modelul matematic

- Modelul ales de noi începe de la orice fișier audio de tip .wav.

- Fișierul audio poate să fie interpretat ca o amplitudine în funcție de timp. Astfel, putem să ne imaginăm că avem o funcție de tipul $f(t)$, unde t este timpul și $f(t)$ este amplitudinea sunetului la acel moment.

- Ca să interpretăm funcția cu acuratețe mai mare, vom aplica Short-Time Fourier Transform, care ne va permite să vedem frecvențele care compun sunetul.

- Astfel o sa ajungem cu o functie $F(\omega, t)$ cu care o sa putem sa modelăm o spectrogramă și pe care o să o putem interpreta ca o matrice.

- Vom aplica SVD (Singular Value Decomposition) pe matricea $F(\omega, t)$, folosind factorizarea QR.

- Am aplicat SVD, vom obține matricea $F(\omega, t) \approx U\Sigma V^T$, unde U și V sunt matricele ortogonale și Σ este o matrice diagonală cu valorile pozitive.

- Inmulțim cele 3 matrici, obținem o aproximare la matricea originală $G(\omega, t)$.

- Cu matricea $G(\omega, t)$ vom calcula Short-Time Fourier Transform inversă, care ne va permite să obținem o nouă funcție $g(t)$, aproximarea funcției originale $f(t)$.

- Cu aproximarea $g(t)$ vom putea să obținem un nou fișier audio .wav, care este o aproximare a fișierului original.

$$\begin{cases} F(\tau,\omega) = \Sigma_{t=0}^{N-1} f(t) w(t-\tau) e^{-i\omega t} \\ g(t) = \frac{1}{w(t-\tau)} \frac{1}{2\pi} \Sigma_{\omega=0}^{N-1} G(\tau,\omega) e^{+i\omega t} \\ F(i,j) = A_{i,j}, A \in \mathbb{M}_{N,M}(\mathbb{C}) \\ G(i,j) = B_{i,j}, B \in \mathbb{M}_{N,M}(\mathbb{C}) \\ A \approx U \Sigma V^T = B \\ w(n) = \frac{1}{2}[1 + cos(\frac{2\pi n}{L})], w : \mathbb{R} \to \mathbb{R} \end{cases} \tag{1}$$

- $w(n)$ este window function Hann
- A și B sunt matricele de amplitudine și frecvență
- B se calculează utilizând metoda SVD și factorizarea Gram-Schmidt

2 Discretizarea domeniului

Domeniul este reprezentat de catre durata melodiei care este înmultită cu sampling rate-ul de 44100 Hz.

$$N = l \times h \tag{2}$$

- l este durata melodiei în secunde
- h este sampling rate-ul default de 44100 Hz

3 Transformata Fourier

Teoretic vorbind, Short-Time Fourier Transform este:

$$F(\tau,\omega) = \int_{-\inf}^{+\inf} f(t) w(t-\tau) e^{-i\omega t} dt \tag{3}$$

dar funcția $f(t)$ este discretizată, deci vom folosi suma discretă ca să calculăm $F(\tau,\omega)$.

$$F(\tau,\omega) = \Sigma_{t=0}^{N-1} f(t) w(t-\tau) e^{-i\omega t} \tag{4}$$

Cu funcția $F(\tau,\omega)$ vom calcula spectograma originală pe care o vom nota cu S_o

$$S_o = \begin{bmatrix} F(\tau_1, \omega_M)^2 & F(\tau_2, \omega_M)^2 & ... & F(\tau_N, \omega_M)^2 \\ ... & ... & ... & ... \\ F(\tau_1, \omega_2)^2 & F(\tau_2, \omega_2)^2 & ... & F(\tau_N, \omega_2)^2 \\ F(\tau_1, \omega_1)^2 & F(\tau_2, \omega_1)^2 & ... & F(\tau_N, \omega_1)^2 \end{bmatrix} \tag{5}$$

4 Sistemului Liniar

In același timp, vom avea matricea:

$$A = \begin{bmatrix} F(\tau_1, \omega_1) & F(\tau_1, \omega_2) & ... & F(\tau_1, \omega_M) \\ F(\tau_2, \omega_1) & F(\tau_2, \omega_2) & ... & F(\tau_2, \omega_M) \\ ... & ... & ... & ... \\ F(\tau_N, \omega_1) & F(\tau_N, \omega_2) & ... & F(\tau_N, \omega_M) \end{bmatrix} \tag{6}$$

Teoretic vorbind, factorizarea SVD este:

$$A = U \Sigma V^T \tag{7}$$

dar o să avem eroare cauzată de floating point, deci vom obține o matrice aproximativă B care este diferită de cea de la care am pornit.

$$B = U \Sigma V^T \approx A \tag{8}$$

Putem să creăm o funcție $G(\tau, \omega)$ care primeste valorile din matricea B și care este definită ca:

$$G(\tau, \omega) = \begin{cases} B_{i,j} & \text{dacă } i \leq N \text{ și } j \leq M \\ 0 & \text{altfel} \end{cases} \tag{9}$$

Utilizând noua funcție, putem crea o nouă spectogramă care o să ne arate eroarea de aproximare pe care o avem:

$$S_e = \begin{bmatrix} |F(\tau_1, \omega_M)^2 - G(\tau_1, \omega_M)^2| & |F(\tau_2, \omega_M)^2 - G(\tau_2, \omega_M)^2| & ... & |F(\tau_N, \omega_M)^2 - G(\tau_N, \omega_M)^2| \\ ... & ... & ... & ... \\ |F(\tau_1, \omega_2)^2 - G(\tau_1, \omega_2)^2| & |F(\tau_2, \omega_2)^2 - G(\tau_2, \omega_2)^2| & ... & |F(\tau_N, \omega_2)^2 - G(\tau_N, \omega_2)^2| \\ |F(\tau_1, \omega_1)^2 - G(\tau_1, \omega_1)^2| & |F(\tau_2, \omega_M)^2 - G(\tau_2, \omega_M)^2| & ... & |F(\tau_N, \omega_1)^2 - G(\tau_N, \omega_1)^2| \end{bmatrix} \tag{10}$$

5 Transformata Fourier Inversă

Calculam aproximarea funcției originale $f(t)$ folosind Short-Time Fourier Transform inversă.

Teoretic vorbind, Short-Time Fourier Transform inversă este:

$$g(t) = \frac{1}{w(t-\tau)} \frac{1}{2\pi} \int_{-\inf}^{+\inf} G(\tau, \omega) e^{+i\omega t} d\omega \tag{11}$$

dar funcția $G(\tau, \omega)$ este discretizată, deci vom folosi suma discretă ca să calculăm $g(t)$.

$$g(t) = \frac{1}{w(t-\tau)} \frac{1}{2\pi} \Sigma_{\omega=0}^{N-1} G(\tau, \omega) e^{+i\omega t} \tag{12}$$

Cu funcția $g(t)$ vom creea o nouă funcție care reprezintă eroarea la un anumit timp f_{ea} și funcția f_{cea} care reprezină eroarea cumulată a aproximării.

$$f_{cea} = \Sigma_{t=0}^{N-1} |f(t) - g(t)| \tag{13}$$

$$f_{ea} = |f(t) - g(t)| \tag{14}$$