

Métodos Numéricos Avanzados

Escala Musical

María de la Puerta Echeverría (50009), Teresa Fontanella De Santis (52455) y Tomás Mehdi (51014)
Grupo 2 - Trabajo Práctico 2 - Instituto Tecnológico de Buenos Aires (ITBA)

Resumen—En el presente artículo se describen dos programas: a) uno que, dado un archivo de sonido, indica el tono fundamental del mismo y crea una partitura asociada; y b) otro que, dado una partitura, crea el sonido correspondiente. Ambos están implementados en Matlab y utilizan la escala de temperamento igual.

Palabras claves—Temperamento igual, FFT, IFFT, tono, semitono, partitura, escala musical.

1 INTRODUCCIÓN

Existe una variedad de escalas musicales que tienen diferentes orígenes históricos. Una de ellas es la escala conocida como de *Temperamento Igual*. Ésta divide la escala de frecuencias musicales en fragmentos compuestos por doce niveles diferentes¹, donde los símbolos \flat (*bemol* en castellano) y \sharp (*sostenido* en castellano) representan los *semitonos*. Asimismo, en la música resulta relevante identificar la nota fundamental de una canción (o *tono*)² y su partitura, y realizar el proceso inverso (dada una partitura, crear el sonido asociado). En este informe se presenta el diseño de dos programas en Matlab que realizan dichas funciones, utilizando la escala de Temperamento igual.

En la siguiente sección, Metodología Utilizada, se describe la misma, así como las pruebas realizadas. En la última sección se muestran los resultados obtenidos y se extraen conclusiones.

2 METODOLOGÍA UTILIZADA

Para la realización de este trabajo, se toman en cuenta las siguientes etapas:

2.1 Modelo empleado

Cada grupo de doce notas implica duplicar la frecuencia fundamental de la escala (es decir, pasar de una octava a la siguiente) y, en el temperamento igual, la relación de frecuencias entre notas sucesivas es siempre la misma: El transmisor envía un dato s_k cada T segundos, donde s_0 es enviado en $t = 0$, s_1 en $t = T$, s_2 en $t = 2T$, ... La modificación de los datos por el canal se presenta

por la denominada *respuesta al impulso del canal* $\{h_k\}_{k=0}^L$, donde L es la longitud de la respuesta al impulso.

Además de ser modificados por el canal, los datos son afectados por ruido blanco Gaussiano aditivo $N_k \sim N(0, \sigma)$.

Teniendo todo esto en cuenta, cada T segundos el receptor observa:

$$r_n = \sum_{k=0}^{L-1} h_k s_{n-k} + N_n \quad (1)$$

También podemos expresar la Ec. (1) en forma matricial como

$$\vec{r} = H\vec{s} + \vec{N} \quad (2)$$

donde

$$\vec{r} = \begin{pmatrix} r_0 \\ r_1 \\ \dots \\ r_M \end{pmatrix}, \vec{s} = \begin{pmatrix} s_0 \\ s_1 \\ \dots \\ s_{M-1} \end{pmatrix}, \vec{N} = \begin{pmatrix} N_0 \\ N_1 \\ \dots \\ N_{M-1} \end{pmatrix} \quad (3)$$

$$H = \begin{pmatrix} h_0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ h_1 & h_0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ \ddots & \ddots & \ddots & \ddots & 0 & 0 & \dots & 0 & 0 \\ h_{L-1} & h_{L-2} & \dots & h_1 & h_0 & 0 & \dots & 0 & 0 \\ 0 & h_{L-1} & \dots & h_2 & h_1 & h_0 & \dots & 0 & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & h_1 & h_0 \end{pmatrix} \in \mathbb{R}^{M \times M}. \quad (4)$$

y $M \geq L$.

Otra forma equivalente es la siguiente:

$$\vec{r} = S\vec{h} + \vec{N} \quad (5)$$

donde \vec{r} y \vec{N} son como antes y

1. Que, en la notación latina son: do, do \sharp , re, mi \flat , mi, fa, fa \sharp , sol, la \flat , la, si \flat , si.

2. Estas actividades pertenecen al llamado *Ear training*, en inglés.

$$\vec{h} = \begin{pmatrix} h_0 \\ h_1 \\ \dots \\ h_L \end{pmatrix}, \quad (6)$$

$$S = \begin{pmatrix} s_0 & 0 & \dots & \dots & 0 \\ s_1 & s_0 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 \\ s_{L-2} & s_{L-3} & \dots & s_0 & 0 \\ s_{L-1} & s_{L-2} & \dots & s_1 & s_0 \\ s_L & s_{L-1} & \dots & s_2 & s_1 \\ \vdots & s_{M-1} & s_{M-2} & \dots & s_{M-L} \end{pmatrix} \in \mathbb{R}^{M \times L}. \quad (7)$$

Lo que nos interesa es recuperar correctamente la señal enviada \vec{s} a partir de la señal recibida \vec{r} . Esto no sería tan difícil si se conociesen L y los $\{h_k\}$. En efecto, tomando $M = L$ en la Ec. (2), podemos buscar \vec{s} que minimice:

$$\|H\vec{s} - \vec{r}\|_2^2. \quad (8)$$

El problema es, sin embargo, que en general ni L ni $\{h_k\}$ son conocidos: estimarlos corresponde al problema de estimación del canal. Si L es conocido, una forma de estimar el canal es mediante la Ec. (5) y el método de cuadrados mínimos. En efecto, tómese $M > L$ y envíese una señal conocida por el receptor $\vec{s} \in \mathbb{R}^M$ (denominada secuencia de entrenamiento). Luego, se estima el canal planteando el siguiente problema de cuadrados mínimos: encontrar $\vec{h} \in \mathbb{R}^L$ que minimice

$$\|S\vec{h} - \vec{r}\|_2^2. \quad (9)$$

2.2 Pruebas realizadas

Para efectuar las pruebas correspondientes, se considera utilizar un $L = 1, 10, 30, 50^1$, aplicando un ruido gaussiano $\sigma = 1$. Se considera este valor por ser lo suficientemente chico para no distorsionar demasiado los datos a transmitir. También se prueba con diferentes longitudes de la cadena de entrenamiento $M = 32, 512, 1024$.

Para analizar cómo se estima en "condiciones óptimas" (verbigracia: sin ruido, etc.), también se prueba con

1. Resolviendo dudas con el profesor, se ha llegado a la conclusión de que a valores de L mayores a 10, el resultado obtenido no es satisfactorio. Por este motivo, para que la prueba sea exitosa los valores de L deben ser $L = 1, 2, 3, 4, 5$.

$M = 512$ y $\sigma = 0$ (con los valores de L ya mencionados)².

3 RESULTADOS

Para $L = 1$, $M = 32$ y $\sigma = 1$ se observa que la imagen de Lena se recupera de manera confiable a pesar de que el canal era inestable.

Para $L = 2$, $M = 32$ y $\sigma = 1$ la imagen se recupera en forma errónea dadas dos imágenes con distinto ruido.

Para $L = 3$, $M = 32$ y $\sigma = 1$ la imagen se recupera de manera bastante confiable. Dadas dos imágenes con distinto ruido, se obtienen imágenes recuperadas similares.

Para $L = 4$, $M = 32$ y $\sigma = 1$ las imágenes recuperadas son completamente distintas y distorsionadas, habiendo transmitido dos imágenes con ruido similar.

Para $L = 5$, $M = 32$ y $\sigma = 1$ y una imagen transmitida con poco ruido, ésta se recupera de manera confiable. Para una imagen con más ruido, se recupera una imagen errónea.

Para $L = 1$, $M = 64$ y $\sigma = 1$ no se recupera la imagen perfecta, pero es bastante confiable con respecto a la imagen original que contenía mucho ruido.

Para $L = 2$, $M = 64$ y $\sigma = 1$ se nota una variación importante para dos imágenes con poco ruido. Se recupera una imagen confiable y una muy distorsionada.

Para $L = 3$, $M = 64$ y $\sigma = 1$ se da una situación similar al caso anterior.

Para $L = 4$, $M = 64$ y $\sigma = 1$ y tres imágenes de Lena con distinto ruido, se recuperan dos de manera confiable y una distorsionada.

Para $L = 5$, $M = 64$ y $\sigma = 1$ y tres imágenes de Lena con distinto ruido, se recupera una de manera confiable y dos distorsionadas.

Para $L = 1$, $M = 256$ y $\sigma = 1$ se observan resultados muy similares al caso de igual L y $M = 64$.

Para $L = 2$, $M = 256$ y $\sigma = 1$ y tres imágenes con distinto ruido, se obtienen imágenes confiables.

Para $L = 3$, $M = 256$ y $\sigma = 1$ y tres imágenes con distinto ruido, ninguna se recupera de forma correcta.

Para $L = 4$, $M = 256$ y $\sigma = 1$ y tres imágenes con distinto ruido, se obtienen dos imágenes distorsionadas

2. Debido a que, en la implementación, la imagen se transmitía con $L = 30$ (valor bastante distinto a los valores de L usados para recuperar la información), los resultados obtenidos eran dispares e inconexos. Se acordó con el profesor que, en vez de rehacer todo el trabajo, agregar las pruebas realizadas con este nuevo cambio, en el cual la información se transmite con $L = 5$, usando $\sigma = 1$, $M = 512$ y se considera a la hora de recibir los datos $L = 1, 2, 3, 4, 5$.

Fig. 1. Imagen transmitida (arriba) y recibida (abajo) usando $L = 1, M = 512, \sigma = 1$.

Fig. 2. Imagen transmitida (arriba) y recibida (abajo) usando $L = 3, M = 512, \sigma = 1$.

y una confiable.

Para $L = 5, M = 256$ y $\sigma = 1$ se recupera una imagen confiable a partir de una cuyo ruido es mayor a otras dos transmitidas y sus correspondientes no se recuperan con éxito.

Para $L = 1, M = 512$ y $\sigma = 1$ se recuperan dos imágenes defectuosas pero claras, dadas dos fotografías con gran nivel de ruido.

Para $L = 2, M = 512$ y $\sigma = 1$ se transmiten tres imágenes, dos con nivel de ruido similar y una con mayor cantidad. Para la imagen con mayor ruido, se obtiene una buena recuperación, aunque no perfecta y lo mismo sucede con una de las imágenes con menor cantidad de ruido. Con la otra, se obtiene una imagen defectuosa, aunque el nivel de ruido sea similar a la imagen que fue recuperada con éxito.

Para $L = 3, M = 512$ y $\sigma = 1$ y tres imágenes con distinto ruido aunque similar en nivel, dos se recuperan defectuosas y una confiable.

Para $L = 4$ y $L = 5$ con $M = 512$ y $\sigma = 1$ suceden situaciones similares al caso anterior.

Luego de haber descubierto el error mencionado en la sección anterior, se procede a realizar las pruebas pertinentes.

Para $L = 1, M = 512$ y $\sigma = 1$ y una imagen de Lena, ésta se recupera de manera confiable (aunque con tonalidades más blanquecinas que la original) (ver Fig.1).

Para $L = 3, M = 512$ y $\sigma = 1$ y una imagen de Lena, ésta se recupera de una manera mucho más confiable que en el caso anterior (ver Fig.2).

Para $L = 5, M = 512$ y $\sigma = 1$ y una imagen de Lena, ésta se recupera completamente distorsionada.(ver Fig.3)

Fig. 3. Imagen transmitida (arriba) y recibida (abajo) usando $L = 5, M = 512, \sigma = 1$.

4 CONCLUSIONES

Luego de correr el programa con los distintos valores propuestos, se observan resultados muy variados dependiendo del ruido que se le aplica al valor transmitido. Se observan situaciones en las que dadas dos imágenes con valor de ruido similar, se recuperan imágenes completamente distintas. Antes de la corrección efectuada, era imposible afirmar que este método dependiera de que el ruido sea menor para obtener una recuperación exitosa, ya que se observaban situaciones en las que se obtiene una imagen muy distorsionada, habiendo transmitido una imagen con poco ruido. Después, con el último set de pruebas realizadas, se puede afirmar que tomando valores muy altos de L , la imagen no se recupera bien, pero que, para el resto de los valores de L , se recupera de manera aceptable. Esto puede ocurrir debido a que, aún cuando el L estimado sea igual al real, los errores de precisión provocan imágenes distorsionadas (como ocurre con $L = 5$). Sin embargo, se puede concluir que este método puede usarse para la estimación de un canal de comunicaciones, para L chicos.

BIBLIOGRAFÍA

- [1] John Matthews and Kurtis Fink, *Métodos Numéricos con Matlab*, Fecha de publicación: Mayo 2000 — ISBN-10: 8483221810 — ISBN-13: 978-8483221815 — Edición: 3
- [2] Apuntes de la Cátedra.
- [3] Consultas a la Cátedra, a quien agradecemos su gentil colaboración, por su experta ayuda en la preparación de este artículo.

APÉNDICE

- .1 Código y ejecución paso 1**
- .2 Código y ejecución paso 2**
- .3 Implementación QR completo por G-S**

.4 Otros datos representativos de las pruebas

h	h_estimada
0.0592715994972346	1.22138294001092
0.316475471117918	
0.323844578892922	
0.531842458909969	
0.402201195880913	

Tabla 1 : Valores correspondientes a la prueba con $L = 1, M = 512, 1$ de Fig.1

h	h_estimada
0.445541676764217	0.457982382433996
0.101034422381018	0.147135682171331
0.0637079819621977	0.0806792295209017
-0.132967712825480	
0.219679633044226	

Tabla 2 : Valores correspondientes a la prueba con $L = 3, M = 512, 1$ de Fig.2

h	h_estimada
0.0455337140835005	0.0449846995202603
0.108336297046597	0.108106879455770
0.149204486339813	0.149598377126117
0.453840022112420	0.453617070901002
0.607263144894805	0.608308008105862

Tabla 3 : Valores correspondientes a la prueba con $L = 5, M = 512, 1$ de Fig.3