VV570
Numerical Analysis

Project 1 (21/06/2017)

Jing and Manuel — UM-JI

Guidelines

- Read through the whole project before starting
- Write in a complete style (subject, verb, and object)
- Clearly explain the reasoning process
- Write the report in LaTeX

1 Linear prediction of speech

Speech production is the result of an excitation signal generated by the contraction of the lungs when they expel air. It is then modified by resonances when passing through the trachea, the vocal cords, the mouth cavity, as well as various muscles (fig. 1). The excitation signal is either created by the opening and closing of the vocal cords, or by a continuous flow of air.
Introduced in the early 1960s by Fant, *the source-filter model* assumes that the glottis and vocal tract are fully uncoupled. This initial idea was reused to develop the *Linear Predictive* (LP) model for speech production.

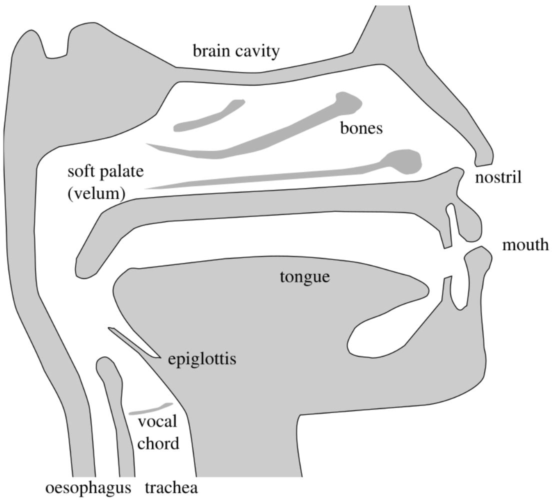

Figure 1: human vocal apparatus

In this model the speech signal is the output $y[n]$ of an *all-pole filter*[1] $1/A(z)$ excited by $x[n]$. Calling $Y(z)$ and $X(z)$ the Z-transform of the speech and excitation, respectively, the model is described by

$$Y(z) = \frac{1}{1 - \sum_{i=0}^{p} a_i z^{-i}} X(z) = \frac{1}{A_p} X(z).$$

Applying the inverse Z-transform to this equation we observe that the speech can be linearly predicted from the previous p samples and some excitation:

$$y[n] = x[n] + \sum_{i=1}^{p} a_i y[n-i].$$

Our goal is to explain as much as possible of $y[n]$ through the a_i, i.e. we look at $x[n]$ as an error, and we strive at rendering it as small and simple as possible. For the sake of clarity we therefore rename $x[n]$ into $e[n]$. The question we want to answer is how to select the a_i such as to minimize the energy $E = \sum_{m=-\infty}^{\infty} e^2[m]$.

1. Show that

$$\sum_{m=-\infty}^{\infty} y[m]y[m-i] = \sum_{i=1}^{p} a_i \sum_{m=-\infty}^{\infty} y[m-i]y[m-i], \quad i = 1, \cdots, p.$$

Since those sums are infinite they cannot be computed, and as such need to be truncated.

[1] A filer whose frequency response function goes infinite at specific frequencies.

This can be achieved by applying the covariance method, which consists in windowing the error

$$E_n = \sum_{m=n}^{n+N-1} \left(y[m] - \sum_{i=1}^{p} a_i y[m-i] \right)^2 .$$

2. Prove that when the error is minimized then

$$\Phi_n(k,0) = \sum_{i=1}^{p} a_i \Phi_n(k,i), \quad \text{with } \Phi_n(k,i) = \sum_{m=n}^{n+N-1} y[m-k]y[m-i].$$

3. Conclude that

$$\begin{pmatrix} \Phi(1,0) \\ \vdots \\ \Phi(p,0) \end{pmatrix} = \underbrace{\begin{pmatrix} \Phi(1,1) & \cdots & \Phi(1,p) \\ \vdots & & \vdots \\ \Phi(p,1) & \cdots & \Phi(p,p) \end{pmatrix}}_{\Phi} \begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix}.$$

Determining the optimal value for the a_i, $1 \le i \le p$, implies inverting the matrix Φ. This can be achieved through Cholesky decomposition.

2 Linear algebra

2.1 Reminders

Let E and F be two vector spaces of finite dimension over $\mathbb{K}$, and u $\mathcal{L}(E,F)$. If $(e_1, \cdots, e_n)$ and $(f_1, \cdots, f_n)$ are the bases of E and F, respectively, then we have $u(e_i) = \sum_{i=1}^{m} a_{i,j} f_i$, and u can be represented by the matrix

$$A = \begin{pmatrix} a_{1,1} & \cdots & \cdots & a_{1,n} \\ \vdots & & & \vdots \\ a_{m,1} & \cdots & \cdots & a_{m,n} \end{pmatrix}.$$

1. Recall the following definitions.
   (a) Hermitian matrix;
   (b) Transpose of A, A^T;
   (c) Trace of A, $\mathrm{Tr}(A)$;
   (d) Eigenvalues and eigenvectors of A;
   (e) Spectrum of A, σ_A;
   (f) Spectral radius of A, $\rho(A)$;
2. Prove that if A is hermitian, then the spectrum of A is a subset of $\mathbb{R}$.
3. Let A and B be two square matrices of order n. Show that $\sigma_{AB} = \sigma_{BA}$ and $\rho(AB) = \rho(BA)$.
4. Recall what a matrix norm is and show that $N_1(A) = \sum_{i,j} |a_{i,j}|$, $N_2(A) = \left(\sum_{i,j} |a_{i,j}|^2\right)^{1/2}$, and $N_\infty(A) = \max\{|a_{i,j}| : 1 \le i,j \le n\}$ are matrix norms.

Similar to how the norm over $\mathcal{L}(E,f)$ was constructed, the norm N_i induces the norm

$$\|A\|_i = \sup_{\substack{x \neq 0 \\ x \in \mathbb{R}^n}} \frac{N_i(Ax)}{N_i(x)} = \sup_{\substack{|x| \le 1 \\ x \in \mathbb{R}^n}} N_i(Ax) = \sup_{\substack{|x| = 1 \\ x \in \mathbb{R}^n}} N_i(Ax)\}.$$

If A is invertible and $\|\cdot\|$ is an induced norm then we define the *condition number* $\kappa(A) = \|A\|\|A^{-1}\|$.

2.2 Linear system of equations

We want to solve the following system of linear equations

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n = b_n \end{cases}$$

1. Show that this system of equation can be rewritten in term of matrices as $AX = B$.
2. Show that the solution X is unique if and only if A is invertible in $\mathbb{K}$.

Let A be invertible and $\|\cdot\|$ be the matrix norm induced by a norm $N(\cdot)$.

3. Prove that if X is the unique solution of the linear system $AX = B \neq 0$, and for a small variation Δ, $X + \Delta X$ is a solution of $(A + \Delta A)(X + \Delta X) = B$, then

$$\frac{N(\Delta X)}{N(X + \Delta X)} \leq \kappa(A)\frac{\|\Delta A\|}{\|A\|}.$$

From this result we observe that the condition number measures the sensitivity of the solution of a linear system. It is therefore a important indicator of the stability of a linear numerical method.

2.3 Cholesky decomposition

A symmetric matrix A over $\mathbb{R}$ is *positive define* if and only if $x^T Ax$ is strictly positive for any $x \neq 0$.

1. Let A be a symmetric matrix. Prove that A is positively defined if and only if its eigenvalues are positive.
2. Let A be a positive define symmetric matrix of order n. Prove the existence of a lower triangular matrix L such that $A = LL^T$. The process of rewriting A is called Cholesky decomposition.
   Hint: proceed by induction on the order of the matrix.
3. Write the pseudocode of an algorithm which returns the Cholesky decomposition of an input matrix A.

3 Tasks

The main goal is to write an **applied math paper** containing all the necessary information to perform linear prediction of speech. Note that beside the actual content the article should feature a title, an abstract and the author's name and affiliation.

3.1 General content

All the information should be organised in a clear and easy to understand fashion **for the reader**. In particular **do not answer the questions one by one as formulated in the preceding sections**. Moreover observe that some information from the previous sections might be useless while some other might me missing. It is important to rightly select the necessary and only the necessary information in order to create a consistent whole.

3.2 Experiments

Generate a sound sample and model it through the LP model. During that stage two strategies can be applied to determine the best choice of a_i as defined in section 1: either window the error and apply the Cholesky decomposition or window the signal. In the later case a different algorithm is to be applied, thoroughly investigate it. Finally, discuss the results and assess when each method performs best.
More information regarding the LP model can be found in [3, 1], while [2] provides insightful results regarding the mathematical side of the problem.

3.3 Remarks

It is recommended to write the code in MATLAB, but C and C++ are also allowed. Although MATLAB provides high-level functions, **algorithms, such as the Cholesky decomposition, must be implemented**. The code of the whole project is to be submitted along with the paper.
Any external resource can be used, however never copy/paste anything and clearly cite any source of information used. Note that similarity checks will be run on the source code. Anybody found cheating or reusing the code or materials of someone else **will be sent to Honor Council**.
Note that although useful information can be found on Wikipedia it is unlikely to be precise and thorough enough to meet the requirements of an applied math paper.

References

[1] Thierry Dutoit and Ferran Marques. *Applied Signal Processing: A MATLAB-Based Proof of Concept*. Springer Publishing Company, Incorporated, 1st edition, 2009.

[2] Martin H. Gutknecht and Marlis Hochbruck. The stability of inversion formulas for toeplitz matrices. *Linear Algebra and Its Applications*, 223-224(Special Issue Honoring Miroslav Fiedles and Vlastimil Ptak):307–324, 1995.

[3] Ian McLoughlin. *Applied Speech and Audio Processing: With Matlab Examples*. Cambridge University Press, New York, NY, USA, 1st edition, 2009.