Eigenanalysis

The Eigenvalue Problem

Let **R** denote the M x M correlation matrix of a WSS process represented by the M x 1 observation (realization) $\mathbf{u}(n)$. We wish to find the M x 1 vector **q** (called the eigenvector) that satisfies

$$\mathbf{Rq} = \lambda\mathbf{q}$$

for some λ (called the eigenvalue). We can rewrite the above as

$$(\mathbf{R} - \lambda\mathbf{I})\mathbf{q} = \mathbf{0}$$

where **I** is the M x M identity matrix and **0** is the M x 1 null vector. The matrix $(\mathbf{R} - \lambda\mathbf{I})$ has to be singular. Therefore we have a non-trivial (non-zero) solution **q** if and only if

$$\det(\mathbf{R} - \lambda\mathbf{I}) = 0 \text{. (characteristic equation)}$$

When the determinant is expanded, we have an order M polynomial in λ. In general we have M roots of this polynomial which lead to M solutions for **q**.

Note that if $\mathbf{q}_i$ is an eigenvector associated with eigenvalue λ_i, then so is $a\mathbf{q}_i$, $a \neq 0$, i.e.

$$\mathbf{R}a\mathbf{q} = \lambda a\mathbf{q}$$

Therefore we typically choose **q** so that it is normalized to have length 1

$$\mathbf{q}_i^H \mathbf{q}_i = 0, \quad i = 1, \ldots, M$$

Example 1: Consider the M x M correlation matrix of a white-noise process where

$$\begin{aligned} r(l) &= E[u(n)u^*(n-l)] \\ &= \sigma^2 \delta(l) \end{aligned}$$

and

$$\mathbf{R} = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma^2 \end{bmatrix}.$$

Here we have a single eigenvalue with multiplicity M

$$\lambda_i = \sigma^2, \; i = 1, \ldots, M \; .$$

Any M x 1 vector **q** is a valid eigenvector, i.e.

$$\mathbf{Rq} = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \ddots & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \sigma^2 \end{bmatrix} \begin{bmatrix} q_1 \\ \vdots \\ q_1 \end{bmatrix}$$

$$= \begin{bmatrix} \sigma^2 q_1 \\ \vdots \\ \sigma^2 q_1 \end{bmatrix} = \sigma^2 \begin{bmatrix} q_1 \\ \vdots \\ q_1 \end{bmatrix} = \lambda \mathbf{q}$$

Note that the power spectrum of such process might look like

Figure: PSD of white noise zero mean, unit variance, Gaussian.

Thus we see again that the eigenvalues of the correlation matrix can be approximated as uniformly spaced samples of the power spectrum.

Properties of Eigenvalues and Eigenvectors

Property 1. Let $\mathbf{q}_1, \ldots, \mathbf{q}_M$ be the eigenvectors corresponding to the distinct eigenvalues $\lambda_1, \ldots, \lambda_M$ of the $M \times M$ correlation matrix **R**, respectively. The eigenvectors $\mathbf{q}_1, \ldots, \mathbf{q}_M$ are *linearly independent*. (A set of vectors is linearly independent if

$$\sum_{i=1}^{M} \nu_i \mathbf{q}_i = \mathbf{0}$$

only when $\nu_1, \ldots, \nu_M$ are all zero.)

Proof: (See text)

The linearly independent vectors $\mathbf{q}_1, \ldots, \mathbf{q}_M$ can serve as a basis for the representation of an arbitrary $M \times 1$ vector **w**. In particular we may express **w** as

$$\mathbf{w} = \sum_{i=1}^{M} \nu_i \boldsymbol{\theta}_i$$

where $\nu_1, \ldots, \nu_M$ are constants.

Property 2. Let $\lambda_1, \ldots, \lambda_M$ be the eigenvalues of the $M \times M$ correlation matrix **R**. Then all the eigenvalues are real and nonnegative.

Proof: Consider the i^{th} eigenvalue/eigenvector in the equation

$$\mathbf{Rq}_i = \lambda_i \boldsymbol{\theta}_i .$$

Premultiplying we have

$$\mathbf{q}_i^H \mathbf{Rq}_i = \lambda_i \boldsymbol{\theta}_i^H \boldsymbol{\theta}_i$$

or

$$\lambda_i = \frac{\mathbf{q}_i^H \mathbf{R}\mathbf{q}_i}{\mathbf{q}_i^H \mathbf{q}_i}$$

since $\mathbf{q}_i^H \mathbf{q}_i > 0$. Since **R** is positive semi-definite we have

$$\lambda_i = \frac{\mathbf{q}_i^H \mathbf{R}\mathbf{q}_i}{\mathbf{q}_i^H \mathbf{q}_i} \geq 0$$

for all *i*. Since **R** is usually positive definite, we typically have

$$\lambda_i > 0 \; .$$

We call the ratio of $\mathbf{q}_i^H \mathbf{R}\mathbf{q}_i$ (Hermitian form) to $\mathbf{q}_i^H \mathbf{q}_i$ (inner product) the Rayleigh quotient of the vector $\mathbf{q}_i$.

Property 3. Let $\mathbf{q}_1, K \; , \mathbf{q}_M$ be the eigenvectors corresponding to the distinct eigenvalues $\lambda_1, K \; , \lambda_M$ of the $M \times M$ correlation matrix **R**, respectively. Then the eigenvectors $\mathbf{q}_1, K \; , \mathbf{q}_M$ are orthogonal to each other, i.e.

$$\mathbf{q}_i^H \mathbf{q}_j = 0, \;\; \iota \neq \varphi .$$

Proof: (See text)

Property 4. Let $\mathbf{q}_1, K \; , \mathbf{q}_M$ be the orthonormal eigenvectors, i.e.

$$\mathbf{q}_i^H \mathbf{q}_j = \begin{cases} 1, & \iota = \varphi \\ 0, & \iota \neq \varphi \end{cases}$$

corresponding to the distinct eigenvalues $\lambda_1, K \; , \lambda_M$ of the $M \times M$ correlation matrix **R**, respectively. Define

$$\mathbf{Q} = [\boldsymbol{\theta}_1 \;|\; \boldsymbol{\theta}_2 \;|\; \Lambda \;|\; \boldsymbol{\theta}_M]$$

and

$$\Lambda = \text{diag}(\lambda_1, K \; , \lambda_M) \; .$$

Then **R** may be "diagonalized" as follows

$$\mathbf{Q}^H \mathbf{R}\mathbf{Q} = \Lambda \; .$$

Proof: Consider the i^{th} eigenvalue/eigenvector in the equation

$$\mathbf{R}\mathbf{q}_i = \lambda_\iota \boldsymbol{\theta}_\iota \; .$$

We may rewrite the set of *M* such equations as

$$\mathbf{R}\mathbf{Q} = \boldsymbol{\Theta}\Lambda \; .$$

Since the eigenvectors are orthonormal we have

$$\mathbf{Q}^H\mathbf{Q} = \mathbf{I}$$

or that **Q** is unitary

$$\mathbf{Q}^{-1} = \mathbf{\Theta}^H \ .$$

Thus we have

$$\mathbf{Q}^H\mathbf{R}\mathbf{Q} = \Lambda \ .$$

We may also express **R** as

$$\begin{aligned}\mathbf{R} &= \mathbf{\Theta}\Lambda\mathbf{\Theta}^H \\ &= \sum_{\iota=1}^{M} \lambda_\iota \boldsymbol{\theta}_\iota \boldsymbol{\theta}_\iota^H \ .\end{aligned}$$

Property 5. The sum of the eigenvalues $\lambda_1, \mathrm{K}\ , \lambda_M$ (not necessarily distinct) of an $M \times M$ matrix **A** is equal to the trace of the matrix, i.e.

$$\sum_{i=1}^{M} \lambda_i = \sum_{i=1}^{M} \mathbf{A}_{ii} \ .$$

Property 6. If the ratio of the largest to smallest eigenvalue (called eigenvalue spread or condition number assuming the spectral norm) of the correlation matrix **R** is large, then **R** is ill-conditioned. We usually use the notation

$$\chi = \lambda_{\max} / \lambda_{\min} \, .$$

Property 7. (Special case of the minimax theorem) Let $\lambda_{\max}$ be the largest eigenvalue of the $M \times M$ correlation matrix **R**. Then

$$\max_{\substack{\mathbf{x} \in X^M \\ \boldsymbol{\xi} \neq \mathbf{0}}} \frac{\boldsymbol{\xi}^H \mathbf{P} \boldsymbol{\xi}}{\boldsymbol{\xi}^H \boldsymbol{\xi}} = \lambda_{\mu\alpha\xi} \ .$$