Eigenanalysis of Autocorrelation Matrices in the Presence of Noncentral and Signal-Dependent Noise

Wei Wang, *Student Member, IEEE,* and Tülay Adali, *Senior Member, IEEE*

***Abstract*—We present approximations for eigenvalues of autocorrelation matrices in the presence of noncentral and signal-dependent noise as a function of eigenvalues of noiseless input. We derive error bounds for the approximations and discuss their properties. The results of the eigenanalysis are applied to the study of first-order polarization mode dispersion for optical systems. The simulation results demonstrate a good match between the approximated and true eigenvalues.**

***Index Terms*—Autocorrelation matrix, eigenanalysis, noncentral and signal-dependent noise.**

I. INTRODUCTION

THE input autocorrelation matrix plays an important role in describing the properties of statistical filters, in particular those that are based on second-order statistics. For example, both the convergence and misadjustment characteristics of the least-mean-squares adaptive filter can be explained through eigenanalysis of its input autocorrelation matrix [1].

In this letter, we derive a useful approximation for the eigenvalues of an autocorrelation matrix in the presence of signal-dependent and noncentral noise. These approximations explain the change in the eigenvalues of the matrix when there is signal-dependent and noncentral noise in the system. For example, such a result is of great interest in optical communication systems, as electronic equalizers are becoming increasingly important in these systems [2]. Most optical communication systems rely on direct detection and use simple amplitude-shift-keying modulation formats, such as return-to-zero (RZ). Thus, the signal is nonzero mean, and the photodetectors, which act as square-law devices, introduce signal-dependent noncentral noise into the detected signal, i.e., the electrical current.

In Section II, we present an eigenanalysis for the autocorrelation matrix of an input with signal-dependent and noncentral noise as a function of the eigenvalues of a noiseless input autocorrelation matrix. In this section, we also perform error analysis and study bounds for the given approximations. In the development, we present an interesting result that a certain class of matrices share the *exact same* eigenvectors. This class of matrices approximates the autocorrelation matrix of a weakly correlated signal. In Section III, the results are applied to the eigenanalysis of the output of an optical channel that experiences polarization mode dispersion (PMD): a major source of distortion in terrestrial fiber links [2]. Section IV presents the summary.

Manuscript received March 22, 2004; revised August 5, 2004. This work was supported in part by the National Science Foundation under Award NSF CCR-0123409. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Steven L. Grant.

The authors are with the Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD 21250 USA (e-mail: wangw1@umbc.edu; adali@umbc.edu).

Digital Object Identifier 10.1109/LSP.2004.840913

II. EIGENANALYSIS OF THE AUTOCORRELATION MATRIX

A. Autocorrelation Matrix in the Presence of Noncentral and Signal-Dependent Noise

In order to show how noncentral and signal-dependent noise affects the autocorrelation matrix, let $v(n)$ denote the noiseless input and $u(n) = v(n) + \omega(n) + \zeta(n)$ the noisy input, where $\omega(n)$ is the additive noise component and $\zeta(n)$ is the signal-dependent noise component, such that $\zeta(n) = g(v(n))h(\omega(n))$, where $g(\cdot)$ and $h(\cdot)$ denote two functions. We define the two autocorrelation matrices $\mathbf{R} = E\{\mathbf{u}(n)\mathbf{u}(n)^T\}$ and $\mathbf{R}_0 = E\{\mathbf{v}(n)\mathbf{v}(n)^T\}$ for the noisy and noiseless input vectors, respectively. The sample vectors are defined such that they contain the last M samples, e.g., $\mathbf{v}(n) = [v(n), v(n-1), \ldots, v(n-M+1)]^T$.

We consider the case where the two autocorrelation matrices are related by the formula

$$\mathbf{R} = \gamma\mathbf{R}_0 + \alpha\mathbf{I} + \beta\mathbf{1}\mathbf{1}^T \tag{1}$$

where $\mathbf{1}$ is an $M \times 1$ vector of all 1s, and γ, α, and β are coefficients that depend on the relationship of the signal and noise components. We assume that $\omega(n)$ is independent and identically distributed (i.i.d.). Three cases that can be represented by (1) are as follows:

1) The noiseless signal $v(n)$ is i.i.d. In this case, $\mathbf{R}_0$ is a diagonal matrix, and $g(v(n))$ and $h(\omega(n))$ can be of any form. The coefficients in (1) for this case are

$$\begin{aligned}\gamma &= 1\\ \alpha &= \sigma_\omega^2 + 2\mu_h(r_{vg} - \mu_v\mu_g) + 2\mu_g(r_{\omega h} - \mu_\omega\mu_h)\\ &\quad + \mu_h^2\sigma_g^2 + \mu_g^2\sigma_h^2 + \sigma_g^2\sigma_h^2\\ \beta &= \mu_\omega^2 + 2\mu_v\mu_\omega + 2\mu_h\mu_v\mu_g + 2\mu_g\mu_\omega\mu_h + \mu_h^2\mu_g^2\end{aligned}$$

where μ_v, μ_ω, μ_g, μ_h, r_{vg}, and $r_{\omega h}$ are expectation values of $v(n)$, $\omega(n)$, $g(v(n))$, $h(\omega(n))$, $v(n)g(v(n))$, and $\omega(n)h(\omega(n))$, respectively, and σ_v^2, σ_ω^2, σ_g^2, and σ_h^2 are variances of $v(n)$, $\omega(n)$, $g(v(n))$, and $h(\omega(n))$, respectively.

2) The signal $v(n)$ is wide-sense stationary (WSS); hence, $\mathbf{R}_0$ is Toeplitz and $g(v(n)) = cv(n)$, where c is a constant. The coefficients in (1) can now be written as

$$\begin{aligned}\gamma &= 2c\mu_h + c^2\mu_h^2\\ \alpha &= \sigma_\omega^2 + 2c\mu_v(r_{\omega h} - \mu_\omega\mu_h) + c^2\sigma_h^2\left(\mu_v^2 + \sigma_v^2\right)\\ \beta &= \mu_\omega^2 + 2\mu_v\mu_\omega + 2c\mu_v\mu_\omega\mu_h.\end{aligned}$$

3) The signal $v(n)$ is WSS and $E\{h(\omega(n))\} = 0$, which implies

$$\begin{aligned}\gamma &= 1\\ \alpha &= \sigma_\omega^2 + 2\mu_g r_{\omega h} + \left(\mu_g^2 + \sigma_g^2\right)\left(\mu_h^2 + \sigma_h^2\right)\\ \beta &= \mu_\omega^2 + 2\mu_v\mu_\omega.\end{aligned}$$

Note that the above three cases cover a wide range of applications.

B. Approximation of Eigenvalues

We assume that the input is persistently exciting, and therefore, the noiseless correlation matrix $\mathbf{R}_0$ is a positive definite matrix. It can then be factored as $\mathbf{R}_0 = \mathbf{Q}\mathbf{\Lambda}_0\mathbf{Q}^T$, where $\mathbf{\Lambda}_0$ is a diagonal matrix with all positive and real eigenvalues $\lambda_1^0 > \lambda_2^0 > \cdots > \lambda_M^0$, and $\mathbf{Q}$ is an orthogonal matrix with columns as the corresponding eigenvectors. Define

$$\mathbf{A} \equiv \mathbf{Q}^T\mathbf{R}\mathbf{Q} = \gamma\mathbf{\Lambda}_0 + \alpha\mathbf{I} + \beta\mathbf{Q}^T\mathbf{1}\mathbf{1}^T\mathbf{Q} \equiv \mathbf{D} + \beta\mathbf{q}\mathbf{q}^T \quad (2)$$

where the diagonal entries of $\mathbf{D}$ are given by $d_i = \gamma\lambda_i^0 + \alpha$, and $\mathbf{q} \equiv \mathbf{Q}^T\mathbf{1} = [q_1, q_2, \ldots, q_M]^T$, i.e., q_i is the sum of the entries of the ith eigenvector. Note that $\mathbf{A}$ and $\mathbf{R}$ are similar matrices; hence, they have the same eigenvalues.

The eigenvalues of $\mathbf{A}$ are given by the roots of the *secular equation* [3]:

$$\det(\mathbf{A} - \lambda\mathbf{I}) \equiv f(\lambda) = 1 + \beta\sum_{i=1}^{M}\frac{q_i^2}{d_i - \lambda}. \quad (3)$$

The characteristics of $f(\lambda)$ are shown in Fig. 1(a) for a first-order autoregressive (AR) process, with a coefficient of 0.7. The proximity of zero crossings to the poles in $f(\lambda)$ will allow a useful approximation, as shown next. As observed in the figure, the roots of $f(\lambda)$ are bounded by the d_i's, i.e., the poles.

To obtain an approximation for the maximum eigenvalue, we rewrite (3) as

$$f(\lambda) = 1 + \beta\sum_{i=2}^{M}\frac{q_i^2}{d_i - \lambda} + \frac{\beta q_1^2}{d_1 - \lambda}. \quad (4)$$

Since $f(\lambda)$ has a pole at d_1, the last term dominates the behavior of $f(\lambda)$ for $\lambda \simeq d_1$. The second term, which is the sum of terms associated with the other eigenvalues, is almost constant in the region around $\lambda_{\max}$. Thus, we can write

$$f(\lambda_{\max}) \approx 1 + \beta\sum_{i=2}^{M}\frac{q_i^2}{d_i - d_1} + \frac{\beta q_1^2}{d_1 - \lambda_{\max}}. \quad (5)$$

The solution of $f(\lambda_{\max}) = 0$ yields

$$\tilde{\lambda}_{\max} = \gamma\lambda_{\max}^0 + \alpha + \frac{\beta q_1^2}{1 + \beta\sum_{i=2}^{M}\frac{q_i^2}{(d_i - d_1)}}$$

where we have replaced d_1 in the last term in (5) by its definition given in (2). Note that $\tilde{\lambda}_{\max} < \lambda_{\max}$, where $\lambda_{\max}$ is the *true* maximum eigenvalue of $\mathbf{R}$.

We can make use of a further approximation when the input autocorrelation matrix $\mathbf{R}_0$ is positive. This will be the case when the input is non-negative, which is the typical case in optical communications systems. Then, we can use Perron's theorem [4], which states that the maximum eigenvalue of a positive matrix is simple and positive, and the eigenvector associated with this eigenvalue has all positive elements. This implies that all other eigenvectors of $\mathbf{R}_0$ possess positive and negative entries because of the orthogonality condition among the eigenvectors. Hence, it is reasonable to assume that $|q_i| \ll |q_1|$, $i = 2, 3, \ldots, M$. Thus, the maximum eigenvalue can be further approximated as

$$\tilde{\lambda}_{\max} \approx \gamma\lambda_{\max}^0 + \alpha + \beta q_1^2. \quad (6)$$

Also, note that even when matrix $\mathbf{R}_0$ is not positive, β will typically be small, compared to the eigenvalues, as it represents noise contributions. Hence, the approximation given in (6) will, in general, hold for a larger class of matrices. We address the validity of this assumption in Section II-C.

From the characteristics shown in Fig. 1, note that all roots of (3), except the maximum one, are very close to the corresponding poles. This happens when the magnitudes of βq_i^2 are small compared to the distance between eigenvalues. This observation suggests the use of similar arguments to approximate the other eigenvalues of the correlation matrix $\mathbf{R}$. For example, the approximate minimum eigenvalue of $\mathbf{R}$ can be written as

$$\tilde{\lambda}_{\min} = \gamma\lambda_{\min}^0 + \alpha + \frac{\beta q_M^2}{1 + \beta\sum_{i=1}^{M-1}\frac{q_i^2}{(d_i - \lambda_{\min})}} \approx \gamma\lambda_{\min}^0 + \alpha. \quad (7)$$

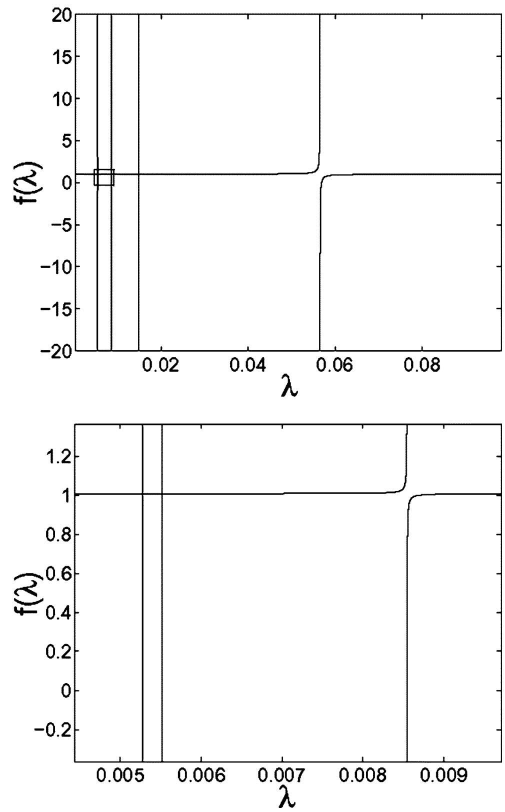

Fig. 1. Example of the characteristics of $f(\lambda)$ for a (Top) 5×5 matrix. (Bottom) close-up of the range shown on the top figure.

C. Error Analysis

In this section, we discuss how close above approximations are to the true eigenvalues. From (4) and (6), we obtain the difference between $\lambda_{\max}$ and $\tilde{\lambda}_{\max}$ as

$$|\Delta\lambda_{\max}| = |\tilde{\lambda}_{\max} - \lambda_{\max}| = \beta q_1^2 \left| \frac{\beta \sum_{i=2}^{M} \frac{q_i^2}{(d_i - \lambda_{\max})}}{1 + \beta \sum_{i=2}^{M} \frac{q_i^2}{(d_i - \lambda_{\max})}} \right|. \tag{8}$$

Substituting

$$\left(1 + \beta \sum_{i=2}^{M} \frac{q_i^2}{d_i - \lambda_{\max}}\right)^{-1} = \frac{(\lambda_{\max} - d_1)}{\beta q_1^2}$$

which is obtained from $f(\lambda) = 0$, into the denominator in (8), we obtain a bound for $|\Delta\lambda_{\max}|$ as

$$|\Delta\lambda_{\max}| = \left|(\lambda_{\max} - d_1)\beta \sum_{i=2}^{M} \frac{q_i^2}{d_i - \lambda_{\max}}\right| \leq \beta \sum_{i=2}^{M} q_i^2 \tag{9}$$

where the last step holds because $\lambda_{\max}$ is greater than all the d_is and closer to d_1 than to any other d_i.

To evaluate the approximation gap in a normalized sense, we divide $\Delta\lambda_{\max}$ by the diagonal entry of $\mathbf{R}_0$, $r_0 = \mu_v^2 + \sigma_v^2$. The reason we choose r_0 is because it is in close scale to the eigenvalues of $\mathbf{R}_0$, according to the Gerschgorin Circle theorem [4], which states that each eigenvalue of a matrix is in at least one of the disks $\{z : |z - r_{ii}| \leq \sum_{j=1, i\neq j}^{M} |r_{ij}|\}$, where r_{ij} is the entry of the matrix $\mathbf{R}_0$. Also, note that r_0 is equal to the arithmetic average of the eigenvalues as $r_0 = (1/M)\sum_{i=1}^{M} \lambda_i$. Hence, we can obtain the error bound of the approximation in (6) as

$$\epsilon_{\max} = \frac{|\Delta\lambda_{\max}|}{r_0} \leq \frac{\beta}{r_0} \sum_{i=2}^{M} q_i^2 \approx \frac{2\mu_w}{\mu_v} \sum_{i=2}^{M} q_i^2. \tag{10}$$

In most cases, $|\mu_w/\mu_v| \ll 1$; hence, a close approximation can be achieved as long as the magnitude of $\sum_{i=2}^{M} q_i^2$ does not become too large. We discuss this quantity later in this section. The error bounds for all the eigenvalues can be obtained in a similar way.

For the approximation of the minimum eigenvalue in (7), a tighter bound can be obtained similarly as

$$|\Delta\lambda_{\min}| = \left| -\beta q_M^2 + \beta q_M^2 \frac{\beta \sum_{i=1}^{M-1} \frac{q_i^2}{(d_i - \lambda)}}{1 + \beta \sum_{i=1}^{M-1} \frac{q_i^2}{(d_i - \lambda)}} \right| \leq \beta q_M^2 \tag{11}$$

since $\sum_{i=1}^{M-1} q_i^2/(d_i - \lambda)$ is positive.

From the error bound in (10), we can see that good approximation to the true maximum eigenvalue is achieved when the magnitudes of q_i^2, $i = 2, 3, \ldots, M$ are small. For positive matrices, the intuition leading to the result is explained using Perron's theorem. However, we need to check whether these quantities are bounded for the two limiting cases. First, this assumption holds very well when the input is highly correlated. In this case, all entries of the correlation matrix are approximately equal (matrix is close to being rank one). Assuming that they are normalized such that input variance is unity, we can write

$$\sum_{k=1}^{M} x_k \approx \lambda x_i \tag{12}$$

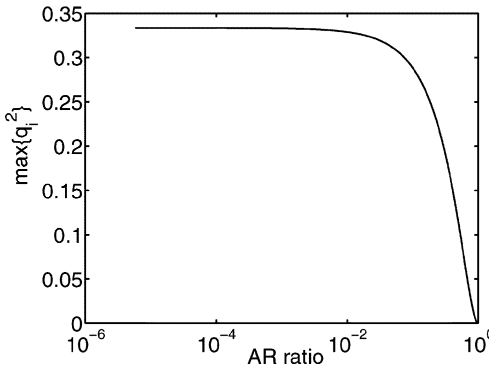

Fig. 2. $\max\{q_i^2\}$, $i = 2, 3, \ldots, M$ for a first-order AR process.

for any $i = 1, 2, \ldots, M$, where $\mathbf{x} = [x_1, x_2, \ldots, x_M]^T$ is the eigenvector associated with λ. Since the left-hand side in (12) is independent of i, the magnitudes of x_is are almost the same, given by $x_i = 1/\sqrt{M}$ for $\mathbf{x}$ normalized such that $\sum_{i=1}^{M} x_i^2 = 1$. Thus, we have $q_i \approx \lambda/\sqrt{M}$. Note that such a matrix has only one dominating eigenvalue, and all the other eigenvalues are close to zero; hence, so are the associated q_i^2s.

On the other hand, when the correlation gets weaker, the magnitudes of q_is increase, thereby reducing the accuracy of the approximation given in (6) and (7). However, the magnitudes of the q_is do not increase indefinitely, and hence, the approximation is still plausible, as for most cases β are small compared to the input eigenvalues. An interesting observation is that when $\mathbf{R}$ approaches a tridiagonal matrix, the magnitudes of q_is converge to constants. In the Appendix, it is shown that all symmetric Toeplitz tridiagonal matrices of the same dimension share common eigenvectors, although their eigenvalues may be different. Fig. 2 shows the maximum q_i^2, $i = 2, 3, \ldots, M$ as a function of the AR coefficient for a first-order AR process. Notice that its magnitude is very small.

III. Application in Optical Communication Systems

Electronic equalization is becoming increasingly important for mitigation of physical impairments in optical communication systems [5]. In particular, it has been shown that it is effective in reducing the signal-to-noise ratio (SNR) penalty due to PMD. In optical systems, bipolar signal transmission formats are seldom used; hence, the transmitted signal is nonzero mean, and the use of photodetectors, which act as square-law devices at the receiver, introduces a noncentral and signal-dependent noise into the received signal, which is the input of the adaptive filters [2].

Let the output of the photodetector be written as

$$u(n) = [s(n) + \eta(n)]^2 \tag{13}$$

where $s(n)$ is the transmitted signal, and $\eta(n)$ is white amplifier noise, distributed $\mathcal{N}(0, \sigma_0^2)$ in the optical domain, prior to detection. Even though the nonlinearities present in the fiber introduce correlation in the amplifier noise, it can be assumed to be white in the band of interest [6]. Let $v(n) \equiv s^2(n)$, $\omega(n) \equiv \eta^2(n)$, and the signal-dependent noise term $\zeta(n) \equiv 2s(n)\eta(n)$. Note that $E\{\eta(n)\} = 0$; hence, it falls into the last category of signals described in Section II-A. The coefficients in (1) for an optical channel are, thus, given by $\gamma = 1$, $\alpha = \sigma_\omega^2 + 4\mu_v\mu_\omega$,

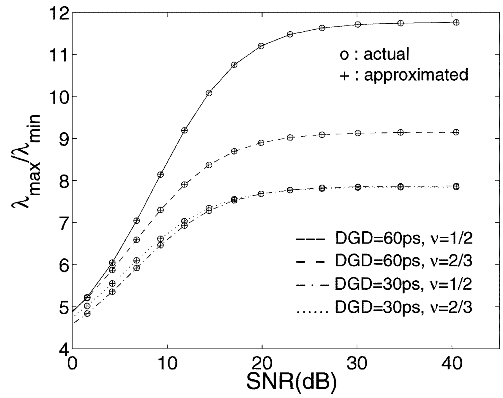

Fig. 3. Eigenvalue spread of $\mathbf{R}$ for a 10-Gbit/s RZ system at the output of a first-order PMD channel with additive and multiplicative noise. The output vector length $M = 5$.

and $\beta = \mu_\omega^2 + 2\mu_v\mu_\omega$. Note that $\mu_\omega = \sigma_0^2$ and the variance of $\omega(n)$ is given by $2\sigma_0^4$.

We can evaluate the eigenvalue spread of autocorrelation matrix using (6) and (7) as

$$\chi(\mathbf{R}) = \frac{\lambda_{\max}}{\lambda_{\min}} \approx \frac{\gamma\lambda_{\max}^0 + \alpha + \beta q_1^2}{\gamma\lambda_{\min}^0 + \alpha}. \tag{14}$$

The results of the analysis can be applied to the output of a first-order PMD channel for optical systems. The effects of first-order PMD can be represented by channel response

$$h_{\mathrm{PMD}}(t) = \nu\delta(t) + (1-\nu)\delta(t-\tau) \tag{15}$$

where ν, a random variable uniformly distributed in [0,1], represents the distribution of power between the two principal states of polarization, and τ is the differential group delay (DGD) value, which is Maxwellian distributed, being the difference in the arrival times of the two polarization states [6].

The first-order PMD channel, as given in (15), is simulated for transmission of 10 Gbit/s RZ Gaussian pulses with 50 ps full width at half maximum. We first compute the eigenvalues of the noiseless input autocorrelation matrix, then apply noise to the system and evaluate the eigenvalue spread by using the proposed method, and finally compare them with the true eigenvalue spreads that are obtained directly from the noisy autocorrelation matrices. A number of channel parameters and noise levels are investigated in the simulations. The SNR is defined as $E\{v^2(n)\}/E\{[\zeta(n) + \omega(n)]^2\}$. As observed in Fig. 3, the given approximations are very close to the actual values for all the cases that are considered. For the example given in this section, $q_1^2 \approx 5$ and q_2^2 take values in the range $[10^{-5}, 10^{-3}]$, which satisfies the assumption that $|q_i| \ll |q_1|, i = 2, 3, \ldots, M$.

IV. Summary

In this paper, we presented an eigenanalysis for the autocorrelation matrices in the presence of noncentral and signal-dependent noise. The results can be used for performance analysis of filters in such cases, e.g., when they are implemented for PMD compensation in optical channels and to improve on their performance [2].

Appendix

Lemma: Two matrices, each with constant main diagonal entries, share the same eigenvectors if the rest of their entries are proportional to each other by a constant factor.

Proof: Let $\mathbf{A}$ be a matrix with constant main diagonal entries, as

$$\mathbf{A} = \begin{pmatrix} a & a_{12} & \dots & a_{1n} \\ a_{21} & a & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a \end{pmatrix} \tag{16}$$

with eigenvalue λ and eigenvector $\mathbf{x}$. Hence

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}. \tag{17}$$

Considering a matrix $\mathbf{B}$ with the following form:

$$\mathbf{B} = \begin{pmatrix} b & a_{12}\beta & \dots & a_{1n}\beta \\ a_{21}\beta & b & \dots & a_{2n}\beta \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}\beta & a_{n2}\beta & \dots & b \end{pmatrix} \tag{18}$$

its eigenvalue μ and eigenvector $\mathbf{v}$ can be obtained by solving

$$(\mathbf{B} - \mu\mathbf{I})\mathbf{v} = \mathbf{0}. \tag{19}$$

The solution of (19) is given by $\mu = b - (a - \lambda)\beta$ and $\mathbf{v} = \mathbf{x}$ since

$$\begin{aligned}(\mathbf{B} - \mu\mathbf{I})\mathbf{x} &= \begin{pmatrix} (a-\lambda)\beta & a_{12}\beta & \dots & a_{1n}\beta \\ a_{21}\beta & (a-\lambda)\beta & \dots & a_{2n}\beta \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}\beta & a_{n2}\beta & \dots & (a-\lambda)\beta \end{pmatrix}\mathbf{x} \\ &= \beta(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} \\ &= \mathbf{0}.\end{aligned}$$

Hence, $\mathbf{B}$ shares the same eigenvectors with $\mathbf{A}$.

Corollary: All symmetric Toeplitz tridiagonal matrices of the same size share common eigenvectors.

Proof: A symmetric Toeplitz tridiagonal matrix is a special case of the matrices described in the lemma in which the diagonal entries are the same, and all the nonzero entries, which are on the two diagonals adjacent to the main diagonal, are constant as well.

References

[1] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 2002.
[2] T. Adalı, W. Wang, and A. O. Lima, "Electronic equalization in optical fiber communications," in *Proc. ICASSP*, Apr. 2003, pp. 497–500.
[3] J. W. Demmel, *Applied Numerical Linear Algebra*. Philadelphia, PA: SIAM, 1997.
[4] R. Bellman, *Introduction to Matrix Analysis*. New York: McGraw-Hill, 1970.
[5] H. Bülow, "Electronic equalization of transmission impairments," in *Proc. OFC*, 2002, pp. 24–25.
[6] G. P. Agrawal, *Fiber-Optic Communication Systems*. New York: Wiley, 1997.