Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 65-76, March 2003

ELA
http://math.technion.ac.il/iic/ela

PROPERTIES OF A COVARIANCE MATRIX WITH AN APPLICATION TO D-OPTIMAL DESIGN*

ZEWEN ZHU†, DANIEL C. COSTER†, AND LEROY B. BEASLEY†

Abstract. In this paper, a covariance matrix of circulant correlation, R, is studied. A pattern of entries in R^{-1} independent of the value ρ of the correlation coefficient is proved based on a recursive relation among the entries of R^{-1}. The D-optimal design for simple linear regression with circulantly correlated observations on $[a, b]$ $(a < b)$ is obtained if even observations are taken and the correlation coefficient is between 0 and 0.5.

Key words. D-optimality, Covariance matrix, Circulant correlation.

AMS subject classifications. 62K05, 15A29.

1. Introduction. D-optimal experimental designs for polynomial regressions on the interval $[-1, 1]$ with uncorrelated observations have been developed; see [4], [9]. However, in the presence of correlations among the observations within each block of the design, these known designs for uncorrelated observations may be inefficient. Atkin and Cheng [1] obtained D-optimal designs for linear and quadratic polynomial regression with a balanced, completely symmetric correlation structure involving a single correlation parameter, ρ. From the results they obtained, we see that D-optimal design in this setting did not always match the known D-optimal designs with uncorrelated observations. Similarly, Kiefer and Wynn [7] studied block designs with a nearest neighbor correlation structure. Properties of the covariance or correlation matrices impacted the optimal designs. In this paper, we consider another correlation structure, that of observations circulantly correlated with the common correlation, and we derive a specific algebraic structure for the inverse of the correlation matrix, which leads to D-optimal simple linear regression design for the observations with the specified correlation structure.

In a statistical linear regression problem with observations $y_1, y_2, \ldots, y_n$ at points $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$, which are in a compact region, the statistical model is

$$y_j = \beta^T \mathbf{f}(\vec{x}_j) + \epsilon_j,$$

where the ϵ_j's are random errors and the variances and covariances among the observations or the random errors are assumed to be

$$cov(y_i,\, y_j) = \begin{cases} \sigma^2 & \text{if } i = j, \\ \sigma^2\rho & \text{if } |i-j| = 1, \text{ or } |i-j| = n-1\ , \\ 0 & \text{otherwise,} \end{cases} \tag{1.1}$$

*Received by the editors on 17 January 2002. Accepted for publication on 25 February 2003. Handling Editor: Michael Neumann.

†Mathematical and Statistical Department, Utah State University, Logan, Utah 84341 ({sl4sv, coster, lbeasley}@cc.usu.edu).

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 65-76, March 2003

i.e., the observations are correlated circulantly. The covariance matrix of the observations will be $\sigma^2 R$, where matrix R is defined as

$$R = \begin{bmatrix} 1 & \rho & 0 & \cdots & 0 & \rho \\ \rho & 1 & \rho & \cdots & 0 & 0 \\ 0 & \rho & 1 & \cdots & 0 & 0 \\ & & \ddots & \ddots & & \\ 0 & 0 & 0 & \cdots & 1 & \rho \\ \rho & 0 & 0 & \cdots & \rho & 1 \end{bmatrix}. \tag{1.2}$$

In regression design and analysis, the covariance matrix of the observations plays a vital role, forming part of information matrix, $M = X^T R^{-1} X$, where

$$X = [\mathbf{f}(\vec{X}_1), \mathbf{f}(\vec{X}_2) ...., \mathbf{f}(\vec{X}_n)]$$

is the design matrix.

This paper derives D-optimal simple linear regression designs with circulant blocks and a correlation structure given by (1.2). Specifically, we show that, in contrast to the uncorrelated case, D-optimality depends not only on the values of the support points, but also on the order of these points. More significantly, the result is shown to hold for any correlation ρ, $0 < \rho < 0.5$, for even block size, and is thus not constrained by nor dependent upon the value of ρ itself. The generality of this D-optimality is a consequence of the pattern of signs of the entries of R^{-1}. Section 2 contains some properties of circulant matrices and a recursive relation among the entries of R^{-1}. The critical result detailing the signs of these entries, for $-0.5 < \rho < 0.5$, is presented in Section 3. Section 4 provides the derivation of the values and order of the support points of a D-optimal design, and examples are given.

2. Preliminary properties. Denote a circulant matrix

$$C = \begin{bmatrix} c_1 & c_2 & c_3 & \cdots & c_{n-1} & c_n \\ c_n & c_1 & c_2 & \cdots & c_{n-2} & c_{n-1} \\ c_{n-1} & c_n & c_1 & \cdots & c_{n-3} & c_{n-1} \\ & & & \ddots & & \\ & & & & \ddots & \\ c_2 & c_3 & c_4 & \cdots & c_n & c_1 \end{bmatrix}$$

by

$$C = cir(c_1, c_2, c_3, \ldots, c_n).$$

Obviously, R is a circulant symmetric matrix. So is R^{-1} if the inverse of R exists. If the correlation coefficient ρ is restricted to $(-0.5, 0.5)$, which is the interval we are interested in, the inverse will exist. We denote it by

$$R^{-1} = cir(v_1, v_2, v_3, \ldots, v_n),$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 65-76, March 2003

ELA
http://math.technion.ac.il/iic/ela

where $v_1, v_2, \ldots, v_n$ are the entries of the first row. For the entries of the first row of R^{-1}, the following relations (2.1), (2.2), and (2.3) hold.

$$v_i = v_{n-i+2}, \tag{2.1}$$

where $i = 1 + \lfloor(\frac{n}{2})\rfloor, \ldots, n$;

$$v_2 = \frac{1 - v_1}{2\rho}; \tag{2.2}$$

and for $2 \le i \le 1 + \lfloor(\frac{n}{2})\rfloor$,

$$v_i = -\frac{v_{(i-1)} + \rho v_{(i-2)}}{\rho}. \tag{2.3}$$

Define matrix L to be

$$L = \begin{bmatrix} 1 & \rho & 0 & \cdots & 0 & 0 \\ \rho & 1 & \rho & \cdots & 0 & 0 \\ 0 & \rho & 1 & \cdots & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & \cdots & 1 & \rho \\ 0 & 0 & 0 & \cdots & \rho & 1 \end{bmatrix}.$$

Assume that D_n is the determinant of L with dimension $n \times n$. L is a Jacobi matrix. From [5], we can find the following relation for the determinants of L matrices,

$$D_n = D_{n-1} - \rho^2 D_{n-2}, \tag{2.4}$$

where D_{n-1} and D_{n-2} denotes the determinants of L's with dimensions $(n-1)\times(n-1)$ and $(n-2) \times (n-2)$, respectively. By matrix operations we find that

$$\det R = D_{n-1} - 2\rho^2 D_{n-2} - 2(-1)^n \rho^n \tag{2.5}$$

and

$$v_1 = \frac{D_{n-1}}{\det R}. \tag{2.6}$$

If we apply the relations (2.2), (2.3), (2.4), (2.5), and (2.6) for different correlation coefficients, ρ, and block size n. We obtain the first $1 + \lfloor \frac{n}{2} \rfloor$ entries of the first row of R^{-1}. Examples are shown in Table 2.1.

From Table 2.1 the following patterns about these entries emerge.

(1) for positive ρ between 0 to 0.5, the odd entries are positive , the even entries are negative.

(2) for negative ρ between -0.5 to 0, all the entries are positive.

These patterns depend only on the sign of ρ. They do not depend on the specific value of ρ and the number of the observations.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 65-76, March 2003

ELA
http://math.technion.ac.il/iic/ela

TABLE 2.1
The first $(1 + \lfloor \frac{n}{2} \rfloor)$ *entries of the first row of* R^{-1}

n(# observations)	ρ (corr. coef.)	$v_{11}, v_{12}, \ldots, v_{1(1+\lfloor \frac{n}{2} \rfloor)}$(entries)
4	0.4	1.8889 , -1.1111 , 0.8889
4	-0.4	1.8889 , 1.1111111 , 0.8889
5	0.4	1.5657 , -0.7071 , 0.2020
5	-0.4	1.7742 , 0.9677 , 0.6452
7	-0.2	1.0911 , 0.2278 , 0.0480 , 0.0120
7	0.2	1.0911 ,-0.2276 , 0.0471 , -0.0078
8	0.2	1.0911 , -0.2277 , 0.0476 , -0.0104 , 0.0041
8	-0.2	1.0911 , 0.2277 , 0.0476 , 0.0104 , 0.0041

THEOREM 2.1. *Assume that a matrix S is invertible and all its row sums equal σ, then its inverse, S^{-1}, has all row sums equal to $1/\sigma$.*

Proof. Assume that vector $\vec{1}$ is a vector with all elements 1. $S\vec{1} = \sigma\vec{1}$ since S has row sums of σ. So $S^{-1}S\vec{1} = \sigma S^{-1}\vec{1}$, that is, $S^{-1}\vec{1} = \frac{1}{\sigma}\vec{1}$, which establishes the theorem. □

Applying Theorem 2.1 to matrix R, we have the following result.

COROLLARY 2.2. *The sum of each row or each column of matrix R^{-1} is $1/1 + 2\rho$.*

3. Proof of Pattern in R^{-1}.

From [11] or [3] we can conclude that the eigenvalues λ_j's, $j = 0, 1, \ldots, n-1$, of R are given by

$$\lambda_j = 1 + 2\rho\cos(\frac{2\pi j}{n}).$$

Further we can represent the entries of R^{-1} in terms of λ_j's as follows

$$v_k = \frac{1}{n}\sum_{j=0}^{n-1} e^{-2\pi ijk/n}\lambda_j^{-1},$$

where v_k's are the entries of the first row of R^{-1} and i is the imaginary unit. However, these existing expressions do not lend themselves to derivation of the D-optimal design(s), in part because the D-criterion function becomes a weighted average of all the eigenvalues, and neither the values nor the order of the support points are evident from this weighted average. In particular, it is not clear whether optimality depends on ρ. Instead, what matters is the pattern of signs of the entries of R^{-1}. This pattern is derived here, using the recursive relations shown in section 2.

The relation (2.4) for the determinants of L and the relation (2.3) for the entries of R^{-1} are two homogeneous second order difference equations. For the general homogeneous second order difference equation, we have the following result from [10].

LEMMA 3.1 (Quinney). *For a homogeneous second order difference equation of the form*

$$y_{n+1} + ay_n + by_{n-1} = 0,$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 65-76, March 2003

ELA
http://math.technion.ac.il/iic/ela

where $n \in \mathbb{N}$, its auxiliary equation

$$r^2 + ar + b = 0$$

has solutions r_1, r_2. Then the general solution form of the homogeneous second order difference equation is

$$\begin{array}{ll} y_n = Ar_1^n + Br_2^n, & \text{if } r_1 \neq r_2 \\ y_n = (A + nB)r_1^n, & \text{if } r_1 = r_2. \end{array}$$

We can use Lemma 3.1 to derive an explicit representation of the entries of the inverse of R.

Lemma 3.2. *The determinant of the matrix L with dimension $n \times n$ is*

$$D_n = A(\frac{1+\sqrt{1-4\rho^2}}{2})^n + B(\frac{1-\sqrt{1-4\rho^2}}{2})^n \ ,$$

where

$$A = \frac{1+\sqrt{1-4\rho^2}}{2\sqrt{1-4\rho^2}} \quad \text{and} \quad B = -\frac{1-\sqrt{1-4\rho^2}}{2\sqrt{1-4\rho^2}} \ .$$

Proof. The auxiliary equation for the homogeneous second order difference equation (2.4) is

$$r^2 - r + \rho^2 = 0$$

Its solutions are

$$r_1 = \frac{1+\sqrt{1-4\rho^2}}{2} \quad \text{and} \quad r_2 = \frac{1-\sqrt{1-4\rho^2}}{2} \ .$$

Now, (2.4) has initial values $D_1 = 1$ and $D_2 = 1 - \rho^2$. Applying Lemma 3.1, we have

$$A(\frac{1+\sqrt{1-4\rho^2}}{2}) + B(\frac{1-\sqrt{1-4\rho^2}}{2}) = 1,$$
$$A(\frac{1+\sqrt{1-4\rho^2}}{2})^2 + B(\frac{1-\sqrt{1-4\rho^2}}{2})^2 = 1 - \rho^2.$$

Solving the above system for A and B and applying Lemma 3.1, the proof is complete. □

Theorem 3.3. *The first $1 + \lfloor \frac{n}{2} \rfloor$ entries of the first row of R^{-1} are given by*

$$v_i = A(\frac{-1+\sqrt{1-4\rho^2}}{2\rho})^i + B(\frac{-1-\sqrt{1-4\rho^2}}{2\rho})^i \ ,$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 65-76, March 2003

ELA
http://math.technion.ac.il/iic/ela

where

$$A = \frac{-(1+v_{11})\sqrt{1-4\rho^2} - (1-4\rho^2)v_{11} - 1}{4\rho\sqrt{1-4\rho^2}} \ ,$$
$$B = \frac{-(1+v_{11})\sqrt{1-4\rho^2} + (1-4\rho^2)v_{11} + 1}{4\rho\sqrt{1-4\rho^2}} \ ,$$

and $i = 1, \ldots, \lfloor \frac{n}{2} \rfloor + 1$.

We can use the same reasoning and procedure as we used in Lemma 3.2 to prove Theorem 3.3. Based on the analytic forms of the entries of R^{-1}, we have the following theorem.

THEOREM 3.4. *Assume that* $v_1, v_2, \ldots, v_{(\lfloor \frac{n}{2} \rfloor+1)}$ *are the first* $\lfloor \frac{n}{2} \rfloor + 1$ *entries of the first row of* R^{-1}. *Then*
(1) *if* $0 < \rho < 0.5$, *the odd entries are nonnegative, the even entries are non-positive;*
(2) *if* $-0.5 < \rho < 0$, *all the entries are nonnegative.*

Proof. To simplify the notation, we define

$$\gamma_1 = \frac{-1+\sqrt{1-4\rho^2}}{2\rho}, \quad \text{and} \quad \gamma_2 = \frac{-1-\sqrt{1-4\rho^2}}{2\rho} \ ,$$

which are the solutions of the auxiliary equation of (2.3). By (2.5), we have

$$D_{n-1} = \frac{r_1^n - r_2^n}{\sqrt{1-4\rho^2}} \quad \text{and} \quad D_{n-2} = \frac{r_1^{n-1} - r_2^{n-1}}{\sqrt{1-4\rho^2}} \ ,$$

where r_1 and r_2 are the solutions of the auxiliary equation of (2.4), which are defined in Lemma 3.2. Thus,

$$\det R = \frac{r_1^n - r_2^n - 2\rho^2(r_1^{n-1} - r_2^{n-1})}{\sqrt{1-4\rho^2}} - 2(-1)^n\rho^n = r_1^n + r_2^n - 2(-1)^n\rho^n$$

and

$$v_1 = \frac{D_{n-1}}{\det R} = \frac{\frac{r_1^n - r_2^n}{\sqrt{1-4\rho^2}}}{r_1^n + r_2^n - 2(-1)^n\rho^n} \ .$$

Since $\gamma_1 = -r_2/\rho$ and $\gamma_2 = -r_1/\rho$, by Theorem 3.3

$$\begin{aligned} v_i &= \frac{1}{2\sqrt{1-4\rho^2}}[(1+v_{11}\sqrt{1-4\rho^2})\gamma_1^{i-1} + (-1+v_{11}\sqrt{1-4\rho^2})\gamma_2^{i-1}] \\ &= \frac{1}{2\sqrt{1-4\rho^2}}[\frac{2r_1^n - 2(-1)^n\rho^n}{r_1^n + r_2^n - 2(-1)^n\rho^n}\frac{(-1)^{i-1}}{\rho^{i-1}}r_2^{i-1} \\ &+ \frac{-2r_2^n - 2(-1)^n\rho^n}{r_1^n + r_2^n - 2(-1)^n\rho^n}\frac{(-1)^{i-1}}{\rho^{i-1}}r_1^{i-1}] \\ &= C[(r_1^n - (-1)^n\rho^n)r_2^{i-1} + (-r_2^n + (-1)^n\rho^n)r_1^{i-1}] \ , \end{aligned} \tag{3.1}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 65-76, March 2003

ELA
http://math.technion.ac.il/iic/ela

where $C = (-1)^{i-1}/\sqrt{1-4\rho^2}(r_1^n + r_2^n - 2(-1)^n\rho^n)\rho^{i-1}$ and $1 \leq i \leq \lfloor \frac{n}{2} \rfloor + 1$. So for $0 < \rho < 0.5$, $C < 0$ if i is even, $C > 0$ if i is odd; for $-0.5 < \rho < 0$, $C > 0$. If $\Delta = [(r_1^n - (-1)^n\rho^n)r_2^{i-1} + (-r_2^n + (-1)^n\rho^n)r_1^{i-1}]$ is nonnegative, the theorem will be established.

For n even,

$$\Delta = [(r_1^n - \rho^n)r_2^{i-1} + (-r_2^n + \rho^n)r_1^{i-1}]. \tag{3.2}$$

Since $r_2 \leq |\rho| \leq r_1$ for $0 < |\rho| < 0.5$, $r_1^n > \rho^n$ and $r_2^n < \rho^n$. Thus, in (3.2), Δ is nonnegative for $0 < |\rho| < 0.5$.

For n odd,

$$\Delta = [(r_1^n + \rho^n)r_2^{i-1} + (-r_2^n - \rho^n)r_1^{i-1}]. \tag{3.3}$$

For $-0.5 < \rho < 0$, $-\rho^n > 0$. Applying $r_2 \leq |\rho| \leq r_1$ for $0 < |\rho| < 0.5$, we find that Δ is nonnegative for $-0.5 < \rho < 0$. But for $0 < \rho < 0.5$, (3.3) can be represented as

$$\begin{aligned}
\Delta &= [(r_1^n + \rho^n)r_2^{i-1} + (-r_2^n - \rho^n)r_1^{i-1}] \\
&= r_1^{n-(i-1)}\rho^{2(i-1)} + \rho^n r_2^{i-1} - r_2^{n-(i-1)}\rho^{2(i-1)} - \rho^n r_1^{i-1} \\
&= \rho^{2(i-1)} r_1^{(i-1)}[r_1^{n-2(i-1)} - \rho^{n-2(i-1)}] \\
&\quad + \rho^{2(i-1)} r_2^{(i-1)}[-r_2^{n-2(i-1)} + \rho^{n-2(i-1)}]
\end{aligned} \tag{3.4}$$

since $r_1 r_2 = \rho^2$. In (3.4), Δ is nonnegative.

Therefore, Δ is nonnegative for $0 < |\rho| < 0.5$, and the theorem is proved. □

4. D-optimal design with circulantly correlated observations. In this section, the properties developed in the above sections is applied to the D-optimal regression design. In linear regression with correlated observations, the order of the regression points affects the statistical performances [12]. Exact design is considered here. An exact design ξ_n with a size n is a sequence of n trails $x_1, x_2, \ldots, x_n$ for support points or treatment levels/combination. The D-optimality criterion is defined by the criterion function

$$\phi[M(\xi)] = -log\,[det\,M(\xi)].$$

If a design ξ_D minimizes this criterion function, the design ξ_D is called a D-optimal design. Equivalently, we can maximize the determinant of the information matrix $M(\xi)$. The D-optimality is related to the volume of the confidence ellipsoid when the estimates are normally distributed [8]. The volume of the confidence ellipsoid is minimized by a D-optimal design.

THEOREM 4.1. *Consider the linear regression model*

$$y_j = \beta_0 + \sum_{i=1}^{d} \beta_i x_{ij} + \epsilon_j,$$

where $X_j = [1, x_{1j}, x_{2j}, \ldots, , x_{dj}]^T \in \Omega$, $j = 1, 2, \ldots, n$, Ω *is a compact region in* $\mathbb{R}^{d+1}$, *and* y_j *is the observation at point* X_j. *If correlations among errors* ϵ*'s are*

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 65-76, March 2003

ELA
http://math.technion.ac.il/iic/ela

defined in (1.1)*, then all circulant permutations of* $\{X_1, X_2, \ldots, X_n\}$ *produce the same information matrix.*

Proof. Define $\varepsilon = [\epsilon_1, \epsilon_2, \ldots, \epsilon_n]^T$. Then $\text{cov}[\varepsilon] = R$; and define the matrix

$$\mathbf{X} = [X_1,\, X_2,\, \ldots,,\, X_n],$$

which is the transpose of the design matrix. The information matrix for this regression is $M = \mathbf{X}R^{-1}\mathbf{X}^T$. Any circulant permutation of $\{X_1, X_2, \ldots, X_n\}$, e.g.,

$$\{X_{n-m+1}, X_{n-m+2}, \ldots, X_n, X_1, X_2, \ldots, X_{n-m}\}$$

can be obtained by $\mathbf{X}P^m$, where the matrix $P = cir(0,\, 1,\, 0,\, \ldots,\, 0)$. By any circulant permutation of the regression points , the information matrix will be $M_p = \mathbf{X}P^m R^{-1}(\mathbf{X}P^m)^T$. Since $R^{-1} = P^m R^{-1}(P^m)^T$,

$$M_p = (\mathbf{X}P^m)R^{-1}(\mathbf{X}P^m)^T = \mathbf{X}(P^m R^{-1}(P^m)^T)\mathbf{X}^T = \mathbf{X}R^{-1}\mathbf{X}^T = M. \quad \square$$

In simple linear regression on $[-1\,, 1]$ with uncorrelated observations, the D-optimal design can be achieved by taking 50% of observations at 1 and -1, respectively, in any order. For example, if there are 10 observations taken, then five observations are taken at 1, and 5 observations is taken at -1, and D-optimality is achieved. But for regression with correlated observations, the optimal design is related with the order to take the regression points and it is possible to achieved D-optimal design with different regression point set [12]. We obtain the following result about D-optimality for simple linear regression on $[-1\,, 1]$ with circulantly correlated observations.

THEOREM 4.2. *Consider the simple linear regression model*

$$y_j = \beta_0 + \beta_1 x_j + \epsilon_j,$$

where $j = 1,\, \ldots,\, ,\, n$ *and* $x_j \in [-1,\, 1]$. *Assume that the correlations among* y_j*'s are defined by* (1.1) *and that* n *is even, i.e., an even number of observations is taken, and* $0 < \rho < 0.5$. *Then one of the circulant permutations of*

$$\underbrace{\{1,\, -1,\, 1,\, -1,\, \ldots,\, 1,\, -1\}}_{n}$$

is a D-optimal design for this simple linear regression problem.

Proof. Let

$$\vec{1} = [1,\, 1,\, \ldots,\, 1]^T \quad \text{and} \quad \vec{x} = [x_1, x_2, \ldots, x_n]^T.$$

The information matrix of this simple linear regression is

$$M = [\vec{1},\, \vec{x}]^T R^{-1}[\vec{1},\, \vec{x}] = \begin{bmatrix} \vec{1}^T R^{-1}\vec{1} & \vec{1}^T R^{-1}\vec{x} \\ \vec{x}^T R^{-1}\vec{1} & \vec{x}^T R^{-1}\vec{x} \end{bmatrix}. \tag{4.1}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 65-76, March 2003

ELA
http://math.technion.ac.il/iic/ela

The determinant of M is

$$\begin{aligned}\det M &= \vec{1}^T R^{-1}\vec{1}\vec{x}^T R^{-1}\vec{x} - (\vec{1}^T R^{-1}\vec{x})^2\\ &= \vec{x}^T(\vec{1}^T R^{-1}\vec{1}R^{-1})\vec{x} - \vec{x}^T(V^{-1}\vec{1}\vec{1}^T R^{-1})\vec{x}\\ &= \vec{x}^T(\vec{1}^T R^{-1}\vec{1}R^{-1} - R^{-1}\vec{1}\vec{1}^T R^{-1})\vec{x}. \end{aligned} \tag{4.2}$$

By Corollary 2.2, (4.2) is simplified to

$$\det M = \frac{n}{1+2\rho}\vec{x}^T R^{-1}\vec{x} - \frac{1}{(1+2\rho)^2}(\sum_{i=0}^{n-1} x_i)^2. \tag{4.3}$$

It is easy to see that $\det M$ is a quadratic form of the regression points $x_1, x_2, \ldots, x_n$ and that it is always nonnegative. So it is a convex function of the regression points $x_1, x_2, \ldots, x_n$; see [2]. It follows that the maximum value of $\det M$ exists and it occurs at vertices of the hypercube $[-1,\, 1]^n$. So we have to take 1's or -1's as regression support points to produce the D-optimal design.

We will re-index the entries of R^{-1}. Assume that

$$R^{-1} = cir(v_0,\, v_1,\, \ldots,\, v_{n-1}). \tag{4.4}$$

Consider

$$\begin{aligned}\vec{x}^T R^{-1}\vec{x} = v_0 &\sum_{i=0}^{n-1} x_i x_i + v_1 \sum_{i=0}^{n-1} 2x_i x_{((i+1) \pmod n)}\\ &+ v_2 \sum_{i=0}^{n-1} 2x_i x_{((i+2) \pmod n)} + \cdots + v_{(\frac{n}{2}-1)} \sum_{i=0}^{n-1} 2x_i x_{((i+\frac{n}{2}-1) \pmod n)}\\ &+ v_{(\frac{n}{2})} \sum_{i=0}^{\frac{n}{2}-1} 2x_i x_{((i+\frac{n}{2}) \pmod n)}\,. \end{aligned} \tag{4.5}$$

We can represent (4.5) as

$$\begin{aligned}\vec{x}^T R^{-1}\vec{x} = nv_0 &+ v_1[\sum_{i=0}^{n-1}(x_i + x_{((i+1) \pmod n)})^2 - 2n]\\ &+ v_2[\sum_{i=0}^{n-1}(x_i + x_{((i+2) \pmod n)})^2 - 2n] + \ldots\\ &+ v_{(\frac{n}{2}-1)}[\sum_{i=0}^{n-1}(x_i + x_{((i+\frac{n}{2}-1) \pmod n)})^2 - 2n]\\ &+ v_{(\frac{n}{2})}[\sum_{i=0}^{\frac{n}{2}-1}(x_i + x_{((i+\frac{n}{2}) \pmod n)})^2 - n]. \end{aligned} \tag{4.6}$$

In the determinant of the information matrix, M, equation (4.3), we can maximize the determinant $\det M$ by minimizing $(\sum_{i=0}^{n-1} x_i)^2$ and maximizing $\vec{x}^T R^{-1}\vec{x}$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 65-76, March 2003

ELA
http://math.technion.ac.il/iic/ela

simultaneously. Take one of the circulant permutations consisting of -1 and 1, for instance, $1,-1,1,-1,\ldots,1,-1$ as regression support points $\{x_0, x_1, \ldots, x_n\}$. It is obvious that $(\sum_{i=0}^{n-1} x_i)^2$ is minimized since n is even.

In (4.6), consider one term, the j-th term: $v_j[\sum_{i=0}^{n-1}(x_i+x_{((i+j) \pmod n))})^2-2n]$. If j is an odd number, $v_j \le 0$ by Theorem 3.4, and $(i+j) \pmod n$ is odd if i is even; $(i+j) \pmod n$ is even if i is odd. So $(x_i+x_{((i+j) \pmod n))})^2=0$ and the j-th term is minimized under this arrangement. If j is an even number, $v_j \ge 0$ by Theorem 3.4, and $(i+j) \pmod n$ is even if i is even; $(i+j) \pmod n$ is odd if i is odd. So $(x_i+x_{((i+j) \pmod n))})^2=4$ and the j-th term is maximized under this arrangement. Therefore (4.6) is maximized under this arrangement and $\det M$ is maximized. By Theorem 4.1, the proof is now complete. □

The following is a symbolic example to illustrate Theorem 4.2.

EXAMPLE 4.3. Assume that $n=6$ and R^{-1} consists of $V_0>0$, $V_1<0$, $V_2>0$, $V_3<0$ in the following way

$$R^{-1}=\begin{bmatrix} V_0 & V_1 & V_2 & V_3 & V_2 & V_1 \\ V_1 & V_0 & V_1 & V_2 & V_3 & V_2 \\ V_2 & V_1 & V_0 & V_1 & V_2 & V_3 \\ V_3 & V_2 & V_1 & V_0 & V_1 & V_2 \\ V_2 & V_3 & V_2 & V_1 & V_0 & V_1 \\ V_1 & V_2 & V_3 & V_2 & V_1 & V_0 \end{bmatrix} .$$

Further assume that the regression support points are x_0, x_1, x_2, x_3, x_4 and x_5. So the determinant of information matrix for simple linear regression is

$$\det M=\frac{6}{1+2\rho}\vec{x}^T R^{-1}\vec{x}-\frac{1}{(1+2\rho)^2}(\sum_{i=0}^{5} x_i)^2 .$$

In $\vec{x}^T R^{-1}\vec{x}$, the terms with V_1 as the coefficient are

$$2(x_0x_1+x_1x_2+x_2x_3+x_3x_4+x_4x_5+x_5x_0),$$

which can be written as

$$\sum_{i=0}^{5} 2x_i x_{(i+1)(\pmod 6))} .$$

We have the similar representations for the terms with the coefficients V_2 and V_3, so

$$\begin{aligned}\det M=&\frac{6}{1+2\rho}(V_0\sum_{i=0}^{5} x_i i+V_1\sum_{i=0}^{5} 2x_i x_{(i+1)(\pmod 6))}\\ &+V_2\sum_{i=0}^{5} 2x_i x_{(i+2)(\pmod 6))}+V_3\sum_{i=0}^{2} 2x_i x_{(i+3)(\pmod 6))})\\ &-\frac{1}{(1+2\rho)^2}(\sum_{i=0}^{5} x_i)^2 .\end{aligned}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 65-76, March 2003

ELA
http://math.technion.ac.il/iic/ela

When $x_0 = 1$, $x_1 = -1$, $x_2 = 1$, $x_3 = -1$, $x_4 = 1$ and $x_5 = -1$, $\det M$ is maximized.

From the analysis in this section, we know that the analytic D-optimal design for simple linear regression on interval $[-1, 1]$. A natural question to ask is what the D-optimal design is on a general bounded interval $[a, b]$, where $a < b$. In such general cases, 0 may not be a valid regression support point, or the interval may not be symmetric about 0. The following propositions will show that the D-optimal design with circulant correlated observations is invariant under scaling and shift transformations.

PROPOSITION 4.4. *Define* $f(x_1, x_2, \ldots, x_n) = \det M$*, where* $x_1, x_2, \ldots, x_n$ *are regression support points and* M *is the information matrix defined in* (4.1)*. Then* f *is invariant under the shift transformation* $\vec{z} = \vec{x} + d\vec{1}$*, where* d *is a constant and* $\vec{x} = [x_1, x_2, \ldots, x_n]^T$*. That is*

$$f(x_1, x_2, \ldots, x_n) = f(x_1 + d, x_2 + d, \ldots, x_n + d).$$

Proof. From (4.2) we know that

$$\begin{aligned} f(x_1, x_2, \ldots, x_n) &= \frac{n}{1+2\rho}\vec{x}^T R^{-1}\vec{x} - \frac{1}{(1+2\rho)^2}(\sum_{i=1}^{n} x_i)^2 \\ &= \frac{n}{1+2\rho}\vec{x}^T R^{-1}\vec{x} - (\vec{x}^T R^{-1}\vec{1})^2, \end{aligned} \tag{4.7}$$

so that

$$\begin{aligned} f(x_1 + d, x_2 + d, \ldots, x_n + d) &= \frac{n}{1+2\rho}(\vec{x} + d\vec{1})^T R^{-1}(\vec{x} + d\vec{1}) - ((\vec{x} + d\vec{1})^T R^{-1}\vec{1})^2 \\ &= \frac{n}{1+2\rho}(\vec{x}^T R^{-1}\vec{x} + 2d\vec{x}^T R^{-1}\vec{1} + d^2\vec{1}^T R^{-1}\vec{1} - [(\vec{x}^T R^{-1}\vec{1})^2 \\ &\quad +2d\vec{x}^T R^{-1}\vec{1}\vec{1}^T R^{-1}\vec{1} + d^2(\vec{1}^T R^{-1}\vec{1})^2] \\ &= \frac{n}{1+2\rho}(\vec{x}^T R^{-1}\vec{x} + 2d\vec{x}^T R^{-1}\vec{1} + d^2\vec{1}^T R^{-1}\vec{1} - [(\vec{x}^T R^{-1}\vec{1})^2 \\ &\quad +\frac{n}{1+2\rho}(2d\vec{x}^T R^{-1}\vec{1} + d^2\vec{1}^T R^{-1}\vec{1})] \\ &= \frac{n}{1+2\rho}\vec{x}^T R^{-1}\vec{x} - (\vec{x}^T R^{-1}\vec{1})^2. \end{aligned} \tag{4.8}$$

The equality (4.8) is obtained by the fact that $\vec{1}^T R^{-1}\vec{1} = \frac{n}{1+2\rho}$. Therefore $f(x_1, x_2, \ldots, x_n) = f(x_1 + d, x_2 + d, \ldots, x_n + d)$. This completes the proof. □

PROPOSITION 4.5. *Let the same setting as Proposition 4.4 hold. Consider the scaling transformation* $\vec{z} = \lambda\vec{x}$*. If* $f(x_1, x_2, \ldots, x_n)$ *has maximum value at* $\vec{x}_0$*, then* $f(z_1, z_2, \ldots, z_n)$ *has maximum value at* $\vec{z}_0 = \lambda\vec{x}_0$*.*

Proof. From (4.2) we have that

$$\begin{aligned} f(z_1, z_2, \ldots, z_n) &= f(\lambda x_1, \lambda x_2, \ldots, \lambda x_n) \\ &= \frac{n}{1+2\rho}\lambda\vec{x}^T R^{-1}\lambda\vec{x} - (\lambda\vec{x}^T R^{-1}\vec{1})^2 \\ &= \lambda^2[\frac{n}{1+2\rho}\vec{x}^T R^{-1}\vec{x} - (\vec{x}^T R^{-1}\vec{1})^2] \end{aligned} \tag{4.9}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 10, pp. 65-76, March 2003

ELA
http://math.technion.ac.il/iic/ela

From (4.9), we can see that $f(z_1, z_2, \ldots, z_n) = \lambda^2 f(x_1, x_2, \ldots, x_n)$. If $f(x_1, x_2, \ldots, x_n)$ has maximum value at $\vec{x}_0$, then $f(z_1, z_2, \ldots, z_n)$ has maximum value at $\vec{z}_0 = \lambda\vec{x}_0$ and the maximum value of $f(z_1, z_2, \ldots, z_n)$ is λ^2 times the maximum value of $f(x_1, x_2, \ldots, x_n)$. □

EXAMPLE 4.6. Kerr and Churchill [6] describe a biological experiment using a circulant block structure. They refer to this as a "loop" design and discuss its statistical application and efficiency under certain conditions, although they do not assume a common correlation of ρ between adjacent observations in a block. However, if such a correlation structure were to be assumed, which would be reasonable if "leakage" or "contamination" existed between adjacent experimental units because of small or modest spatial separation, and for a simple linear regression model, then Theorem 4.2 would apply, and a D-optimal design would be available.

REFERENCES

[1] J.E. Atkins and C.-S. Cheng. Optimal Regression Designs In The Presence of Random Block Effects. *Journal of Statistical Planning Inference*, 77:321–335, 1999.

[2] E. Chong and S. Zak. *An Introduction to Optimization*. Wiley, New York, 1996.

[3] P.J. Davis. *Circulant Matrices*. Second Edition, Chelsea, New York, 1994.

[4] G. Elfving. Optimum Allocation in Linear Regression Theory. *Annals of Mathematical Statistics*, 23:255–262, 1952.

[5] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

[6] M.K. Kerr and G.A. Churchill. Experimental Design for Gene Expression Microarrays. *Biostatistics*, 2:183–201, 2001.

[7] J. Kiefer and H.P. Wynn. Optimum Balanced Block and Latin Square Designs for Correlated Observations. *Annals of Statistics*, 9:737–757, 1981.

[8] A. Pázman. *Foundations of Optimum Experimental Design*. D. Reidel, Boston, 1986.

[9] F. Pukelsheim. *Optimal Design of Experiments*. Wiley, New York, 1993.

[10] D. Quinney. *An Introduction to the Numerical Solution of Differential Equations*. Revised Edition, Wiley, New York, 1987.

[11] C.R. Rao and M.B. Rao. *Matrix Algebra and its Applications to Statistics and Econometrics*. World Scientific, Singapore, 1998.

[12] Z. Zhu. *Optimal Experiment Designs with Correlated Observations*. Ph.D. dissertation, Mathematical and Statistical Department, Utah State University, 2003.