● *Original Contribution*

# COMPUTER-ASSISTED DESIGN OF SURFACE COILS USED IN MAGNETIC RESONANCE IMAGING. II. ROTATIONAL DISCRIMINATION NONLINEAR REGRESSION ANALYSIS AND THE DESIGN OF SURFACE COILS

JOHN H. LETCHER

Department of Mathematical and Computer Sciences, University of Tulsa, Tulsa, Oklahoma 74104, USA

For a number of reasons, it is desirable to fabricate coils which, for a known current, shall produce predetermined values of the magnetic field intensity at a number of points within a nuclear magnetic resonance imager.

The calculation of the magnetic field intensity at a set of points involves the integration of the Biot–Savart equation for all components of the segments of conductor which make up the coil. This process in itself is a rather formidable task. When this process is parameterized in terms of coil diameter, coil spacing, etc. the problem is to determine the values of these parameters to match values of magnetic field intensities which are desired. The problem thereby increases in complexity to the point where, by ordinary methods, the problem becomes intractable.

A generalized solution technique has been developed on a digital computer to implement the rotational discrimination nonlinear regression techniques of Faris, Law and Letcher to find the best solution to this problem. The problem is posed by integrating the Biot–Savart equation. This produces algebraic expressions for incorporation into the optimization program which is executed on a computer in a conversational mode.

This technique was employed to specify the dimensions of a rectangular surface coil for the investigation of the whole human spine.

*Keywords:* Surface coils; Nonlinear optimization; Magnetic resonance imaging; Design of MRI coils.

## INTRODUCTION

In the use of nuclear magnetic resonance imaging, it is sometimes desirable to have and use coils other than the whole body and head units customarily supplied by the manufacturer of the imager. Coils are also needed which are specific to the study of (1) the lumbar spine, (2) the thoracic spine, (3) the knee and shoulder, and (4) one to properly study specimens immediately after removal by surgical techniques. In the design of a new coil, the magnetic field distribution is specified for the desired unit, properly scaled to suit the dimensions and characteristics of the specimen under investigation. One of the purposes of the study undertaken herein is to produce a technique allowing the specification of the desired magnetic field intensities at specific points in space and to determine the shape and dimensions of the coil which, once fabricated, will produce the desired magnetic field distribution. It is assumed that the calculated magnetic field gives a direct measure of the sensitivity of the coil when the coil is used as an antenna in a magnetic resonance imager.

The magnetic field produced by a current moving in a collection of line segments of wire is calculable.[2] That is, by integration of the Biot–Savart equation, it is possible to calculate the magnetic field intensity throughout the experimental volume. However, for the purposes of the current study, the problem is posed backwards. What is needed is to start with the defined magnetic field characteristics and to calculate the parameters of shape, rather than to start with the specific dimensions and shape characteristics of a hypothesized coil and by iteration come as close as possible to the desired result. There is a further problem, that for a shape suitably general in nature, it may not be possible to solve the problem exactly.

## ROTATIONAL DISCRIMINATION NONLINEAR REGRESSION ANALYSIS[4,5]

Let us assume that there exists a problem which can be posed according to a set of rules. In the statement of the problem, a set of independent variables are defined and expressed as a column vector, $\hat{X}$, such that the $i^{th}$ independent variable $X_i$ is the $i^{th}$ element in the vector $\hat{X}$; that is $X_i = (\hat{X})_i$. Furthermore, let it be possible to define a set of objective functions $\{f_j\}, j = 1, \ldots, m$, where each $f$ is a function of the independent variables $\hat{X}$; i.e., $f_j = f_j(\hat{X})$.

The problem that is posed is to find a set of values of the elements of $\hat{X}$ so that the calculated value of each $f_j$ is as close (in absolute terms) to a defined set of constants $\{g_j\}, j = 1, \ldots, m$, the goals for each objective function. Now since each $f_j$ may have a different dimensionality with respect to the others, we shall allow the definition of a set of constants $\{w_j\}, j = 1, \ldots, m$ so as to make the calculation error $|f_j - g_j|$ for each $j$ to be numerically comparable with respect to each other.

In terms of the above definitions, the problem is easily stated. Let a single valued dimensionless differentiable function $Q$, called the total objective function, be defined thus:

$$Q \equiv \frac{1}{2} \sum_{j=1}^{m} (f_j(\hat{X}) - g_j)^2 w_j . \quad (11)$$

This implies that in turn each $f_j$ is a real differentiable function over the allowed values of the independent variables, $\hat{X}$. Each term in the above sum is intrinsically positive, so, therefore, is $Q$. Only when each $f_j(\hat{X}) = g_j$ is $Q$ equal to zero.

The expression for each objective function may be extraordinarily complicated. Furthermore, the dependence of each $f_j$ upon each $X_i$ may be highly nonlinear.

The purpose of this discussion is to derive an algorithm suitable for execution on a digital computer that starts with some initial set of values of the independent variables $\hat{X} = \hat{X}_0$. The calculational process shall provide a recipe for the calculation of a search vector $\delta \hat{X}$ defined as

$$\hat{X} = \hat{X}_0 + \delta \hat{X} \quad (12)$$

so that $Q(\hat{X})$ is lower in value than $Q(\hat{X}_0)$. By performing this process again and again using the calculated $\hat{X}$ as $\hat{X}_0$ in the next step (iteration), the computer will find a set $\hat{X}$ that minimizes the calculated value of $Q$.

If this process were a single valued function, $X_L$,

we may calculate the value of $\frac{\partial Q}{\partial X_L}$. Given an initial value as shown as $X_{0L}$, it is our intent to calculate the search vector $\delta X_L$.

If we construct a line tangent to the function, $\frac{\partial Q}{\partial X}\Big|_{X_L = X_{0L}}$, the slope of this curve is simply $\frac{\partial^2 Q}{\partial X^2}\Big|_{X_L = X_{0L}}$. We now calculate $X$ by seeing at what value of $X$ the value of $\frac{\partial Q}{\partial X} = 0$ (the criterion for a minimum in the calculated value of $Q$). We find:

$$\frac{y_2 - y_1}{X_2 - X_1} = slope = \frac{\partial^2 Q}{\partial X^2}\Big|_{X_L} = \frac{0 - \frac{\partial Q}{\partial X}\Big|_{X_L}}{(X_{0L} + \delta X) - X_{0L}} \quad (13)$$

or

$$\frac{\partial^2 Q}{\partial \hat{X}^2}\Big|_{X = X_{0L}} \delta X = -\frac{\partial Q}{\partial X}\Big|_{X = X_{0L}} . \quad (14)$$

This is the well-known Newton–Raphson iteration scheme.

Generalizing the above to many dimensions we state that:

$$\frac{\partial^2 Q}{\partial \hat{X}^2} \delta \hat{X} = -\frac{\partial Q}{\partial \hat{X}} . \quad (15)$$

Let us now return to the definition of Eq. (11) and differentiate it twice with respect to each $X_j$, giving

$$\frac{\partial Q}{\partial \hat{X}_i} = \sum_j (f_j(\hat{X}) - g_j)\left(\frac{\partial f_j(\hat{X})}{\partial \hat{X}_i}\right) w_j \quad (16)$$

$$\frac{\partial^2 Q}{\hat{X}_{ik}^2} = \sum_j \frac{\partial f_j}{\partial X_i} \frac{\partial f_j}{\partial X_k} w_j + (f_j(\hat{X}) - g_j)\frac{\partial^2 f_j}{\partial X_i \partial X_k} w_j. \quad (17)$$

The convenient vector notation is used so that $\left(\frac{\partial Q}{\partial \hat{X}}\right)_i = \frac{\partial Q}{\partial X_i}$, etc.

The latter expression in Eq. (17) is customarily discarded (that is, the statement is made that the sum of the terms is zero). This is called the Gauss–Newton approximation and this shall be used here. This approximation is particularly valid when we are close to a solution.

For convenience, let us define a matrix $\hat{G}$ such that

$$\hat{G}_{ik} = \sum_j \left(\frac{\partial f_j}{\partial \hat{X}_i}\right)\left(\frac{\partial f_j}{\partial \hat{X}_k}\right) \approx \left(\frac{\partial^2 Q}{\partial \hat{X}^2}\right)_{ik} \quad . \quad (18)$$

This is called the approximate Hessian matrix and it has a number of interesting properties. For example, $\hat{G}$ is a symmetric, real, positive semidefinite matrix so that there exists an orthogonal matrix $\hat{U}$ which diagonalizes $\hat{G}$, thus:

$$\hat{U}^T\hat{G}\hat{U} = \hat{G}_D \quad (19)$$

(where $(\hat{G}_D)_{ik} = 0$ if $i \neq k$).

Stating that $\hat{G}$ is positive semidefinite means that each $(\hat{G}_D)_{kk} \geq 0$. That $\hat{U}$ is an orthogonal matrix means that $\hat{U}$ has the property that

$$\hat{U}^T\hat{U} = \hat{U}\hat{U}^T = \hat{1} \quad (20)$$

where $\hat{U}^T$ is the transpose of $\hat{U}$ $((\hat{U}^T)_{ik} = (\hat{U})_{ki})$ and $\hat{1}$ is the unit matrix $((\hat{1})_{ik} = 1$ if and only if $i = k$, 0 otherwise).

If any of the diagonal values of $\hat{G}_D$ are zero, then the matrix $\hat{G}$ has no inverse, $\hat{G}^{-1}$, which would be defined to have the properties

$$\hat{G}^{-1}\hat{G} = \hat{G}\hat{G}^{-1} = \hat{1} \quad . \quad (21)$$

This can be seen by investigation how an inverse may be obtained for a real positive definite matrix, $\hat{S}$. An orthogonal matrix $\hat{U}$ exists that diagonalizes $\hat{S}$, i.e., $\hat{U}^T\hat{S}\hat{U} = \hat{S}_D$. Equivalently,

$$\hat{S} = \hat{U}\hat{S}_D\hat{U}^T \quad . \quad (22)$$

All of the diagonal elements of $\hat{S}_D$ are positive and nonzero (which is the definition of the term positive definite.)

The inverse of $\hat{S}_D$ is readily obtained:

$$(\hat{S}_D^{-1})_{kk} = \frac{1}{(\hat{S}_D)_{kk}} \quad . \quad (23)$$

Therefore,

$$\hat{S}^{-1} = \hat{U}\hat{S}_D^{-1}\hat{U}^T \quad . \quad (24)$$

which is the inverse since

$$\hat{S}^{-1}\hat{S} = \hat{U}\hat{S}_D^{-1}\hat{U}^T\hat{U}\hat{S}_D\hat{U}^T = \hat{1} \quad (25)$$

$$\hat{S}\hat{S}^{-1} = \hat{U}\hat{S}_D\hat{U}^T\hat{U}\hat{S}_D^{-1}\hat{U}^T = \hat{1} \quad . \quad (26)$$

Even though some of the diagonal elements of $\hat{G}_D$ are zero, it is still possible to define a pseudoinverse $\hat{G}^P$ which has the properties

$$\hat{G}\hat{G}^P\hat{G} = \hat{G} \quad (27)$$

$$\hat{G}^P\hat{G}\hat{G}^P = \hat{G}^P \quad (28)$$

$$(\hat{G}\hat{G}^P)^T = \hat{G}\hat{G}^P \quad (29)$$

and

$$(\hat{G}^P\hat{G})^T = \hat{G}^P\hat{G} \quad . \quad (30)$$

The pseudoinverse of $\hat{G}$ is found in a manner similar to that of the inverse $\hat{S}^{-1}$ except that when a diagonal element of $\hat{G}_D$ is zero then

$$(\hat{G}_D^P)_{kk} = 0 \text{ when } (\hat{G}_D)_{kk} = 0 \quad . \quad (31)$$

Therefore, if we define a diagonal matrix $\hat{G}_D^P$ such that

$$(\hat{G}_D^P)_{kk} = \begin{cases} \dfrac{1}{(\hat{G}_D)_{kk}} & \text{iff } (\hat{G}_D)_{kk} \neq 0 \\ 0 & \text{otherwise} \end{cases} \quad (32)$$

Thus,

$$\hat{G}^P = \hat{U}\hat{G}_D^P\hat{U}^T \quad . \quad (33)$$

We may take the matrix $\hat{G}$, calculate its pseudoinverse $\hat{G}^P$ and premultiply Eq. (15) by $\hat{G}^P$ giving

$$\hat{G}^P\hat{G}\delta\hat{X} = -\hat{G}^P\frac{\partial Q}{\partial \hat{X}} \quad . \quad (34)$$

Consider the matrix $\hat{G}^P\hat{G}$ which can be expressed

$$\hat{G}^P\hat{G} = [\hat{U}\hat{G}_D^P\hat{U}^T][\hat{U}\hat{G}_D\hat{U}^T] = \hat{U}\hat{G}_D^P\hat{G}_D\hat{U}^T \quad . \quad (35)$$

The matrix $\hat{G}_D^P\hat{G}_D$ is a diagonal matrix with only the values of 0 and 1 on the diagonal. For each occurrence of a zero on the diagonal of $\hat{G}_D$, there is a zero in $\hat{G}_D^P\hat{G}_D$.

Let us define a diagonal matrix $\hat{Z}_D$ which has a 1 on the diagonal when $\hat{G}_D$ has a zero and a 0 when $\hat{G}_D$ has a nonzero value. So,

$$\hat{G}^P\hat{G} = \hat{U}[\hat{1} - \hat{Z}_D]\hat{U}^T \equiv \hat{1} - \hat{Z} \quad . \quad (36)$$

Equation (34) is now expressed:

$$\delta\hat{X} = -\hat{G}^P \frac{\partial Q}{\partial \hat{X}} + \hat{Z}\delta\hat{X} \ . \tag{37}$$

In the Faris and Law treatments,[4] the term $\hat{Z}\delta\hat{X}$ is discarded in return for the introduction of a non-unity constant $h$ in Eq. (37)

$$\hat{X} = \hat{X}_0 + h\delta\hat{X} \ . \tag{38}$$

Equation (37) is the result that we wanted to obtain. Finally we may choose either to discard $\hat{Z}\delta\hat{X}$ in Eq. (37) and use an arbitrary value of $h$ or set the value of $h = 1$ and use Eq. (38) as an iteration scheme for evaluating $\delta\hat{X}$. That is,

$$\delta x =: -\hat{G}^P \frac{\partial Q}{\partial \hat{X}} + \hat{Z}\delta\hat{X} \tag{39}$$

where the symbols "=:" should be read, "is obtained from". A comparison of these two techniques will be covered elsewhere.

The interpretation of this technique is straightforward. Given the matrix $\hat{G}$ which expressed the dependencies of $\frac{\partial^2 Q}{\partial \hat{X}^2}$ and all of the cross dependencies, a new coordinate system is chosen by coordinate rotation that is picked to remove these cross term dependencies. In this new representation, the process is able to carry out $n$ Newton–Raphson iteration procedures at the same time ($\hat{G}$ is $n \times n$). By coordinate rotation back to the original coordinates, the calculated search vector may be properly applied. This is the reason that this technique is commonly called nonlinear programming by way of coordinate transformation or by rotational discrimination (against undesirable cross terms).

This efficient algorithm has been implemented by the author as a conversational computer package called HULDRA.[8,9] The researcher may pose the calculated values of the desired objective functions in terms of a FORTRAN subroutine. HULDRA then calls this routine to find the best fit solution to the stated problem.

## THE DESIGN OF MRI GRADIENT COILS

To calculate the shape parameters for coils to produce a desired magnetic field distribution, one defines an objective function for each point in space for which the calculated magnetic field is to be specified. Thus, for a set of points, the magnetic fields $\mathbf{B}_n(\mathbf{r})$ is stated, then the calculated magnetic field at point $r_n$, $\mathbf{B}(\mathbf{r}_n)$ allows the definition of objective function $f_n$:

$$f_n = |\mathbf{B}(\mathbf{r}_n) - \mathbf{B}(\mathbf{r}_n)| \ . \tag{40}$$

In problems of this sort, $g_n = 0$ and $w_n = 1$.

The shape of the coil is specified parametrically to the computer calculation. The length of the entire coil is broken into small contiguous segments, each of which is approximated by a straight line. Three arrays are set up to specify for each of the straight line segments: (1) the change in the $x$ coordinate; (2) the change in the $y$ coordinate; and (3) the change in the $z$ coordinate that one experiences as one moves along each segment of the length of the coil. During the course of the calculation, the total magnetic field at a point is calculated by indexing through these three arrays calling the subroutine given as a previous paper of this series.[2] As each call produces a small vector quantity, these may be added as vectors to produce the resulting magnetic field at the desired observation point. When Maxwell or Helmholtz coil arrangements are used, it is convenient to employ two long straight segments (with currents mutually opposing each other). These are defined to connect the two coils.

This procedure allows the use of a parameter in the calculation of the shape arrays during the course of the optimization. One may allow a parameter of shape to be one or several of the objective functions over which the optimization is to be carried out.

## THE DESIGN OF A RECTANGULAR SURFACE COIL

As an application of this technique, it was the intent to calculate the size and shape parameters of a coil for the investigation of the whole human spine. As a first attempt, a rectangular shape was chosen (fed at the middle of the short end and turned to the center frequency of 21.322 mHz). For the purposes of this problem, four straight line segments were used to represent the coil. The length and width of the coil which needed to be calculated are the design parameters.

The objective function chosen for the selection of the design parameters of this surface coil is then given in Eq. (40). Please note that there is no concern with optimization with regard to signal to noise. This is intentional. That the objective function chosen will select the desired parameters for a surface coil is an axiom. That is, the reciprocity theorem is taken as the assumption that this is a valid technique. If it is desired to optimize by also adding additional objective functions, this is possible and the mathematical machinery for doing so is given in the body of this paper. Yet, including the calculated magnetic field, alone, is a complex process.

In the design of a coil, one must know the distance from the coil which is of primary clinical diagnostic importance. The range of depths that the coil is to be effective must be known (the coil to skin distances (both front and back)). This shows the ratio of coil sensitivity of a near distance to that of the greatest depth to be studied using this coil.

A surface coil of height of 0.795 meters and width of 0.131 meters was constructed and its properties are reported elsewhere.[6] The calculations described herein are also useful in correcting the resulting image for variations in antenna sensitivity as a function of depth. Studies are underway to use the reciprocal of the calculated magnetic field as a filter function for clinical images. These results will also be reported in the next paper in this sequence.[6]

# REFERENCES

1. Law, V.J.; Faris, R.H. Nonlinear programming via rotational discrimination presented at the International Symposium on Mathematical Programming, Princeton University, Princeton, NJ; 1967.
2. Letcher, J.H. Computer-assisted design of surface coils used in magnetic resonance imaging. I. The calculation of the magnetic field. *Magn. Reson. Imaging* (in press).
3. Letcher, J.H. *J. Chem. Phys.* **54**:3215; 1971.
4. Letcher, J.H. Huldra, the generalized planning subsystem submitted to the U.S. Department of Commerce; 1973.
5. Letcher, J.H. Huldra, rotational discrimination nonlinear regression system, an optimization tool; 1985.
6. Yamanashi, W.; Lester P.D.; Letcher, J.H. Computer-assisted design of surface coils used in magnetic resonance imaging. III. The design and construction of two long twin axial antennae for imaging of the whole human spine. *Mag. Reson. Imaging* (in press).