Electronic Transactions on Numerical Analysis.
Volume 24, pp. 1-6, 2006.
Copyright © 2006, Kent State University.
ISSN 1068-9613.

ETNA
Kent State University
etna@mcs.kent.edu

ORTHOGONAL LEAST SQUARES SOLUTIONS FOR LINEAR OPERATORS*

EVA ACOSTA†

Abstract. This paper solves the problem of finding, in a least squares sense, the coefficients of a series expansion of a function in terms of a chosen orthogonal basis from the knowledge not of the function itself but from the action of a linear operator upon it. The coefficiens are evaluated by inner product with a set of functions related to the orthogonal basis through the adjoint operator of the linear operator. Examples for both differential operators and integral ones as well as related properties are given.

Key words. orthogonal polynomials, linear operators, gradient operator, Radon transform

AMS subject classifications. 33C90, 33C47,42C05,42C15, 47A05

1. Introduction. Orthogonal polynomials have properties that are very useful for solving mathematical and physical problems. They provide a convenient method for expanding a function in a series of linearly independent terms. Moreover, in many physical problems the expansion coefficients can be interpreted as physical magnitudes with a specific meaning and that is why among all the sets of orthogonal basis in which a function can be expanded, only one is usually chosen. For instance, the Zernike polynomials are a set of orthogonal polynomials that arise in the expansion of a wavefront function for optical systems with circular pupils [1]. Expansion coefficients in Zernike's basis have concrete physical meaning related to the shape of the distortion of a wavefront from an ideal spherical one. Some of them even have proper name as astigmatism, coma, defocus etc.

On the other hand, in some physical experiments the information concerning the function (the coefficients of which we want to determine) is not the function itself but an indirect magnitude or magnitudes that can be represented as differential or integral operators acting upon the function. For instance local slopes (gradient operator) or line integrals (Radon transform) etc. The problem is how to evaluate the expansion coefficients in the chosen orthogonal basis since the action of the linear operators upon the orthogonal polynomials give rise to a new set of functions that, in general, are no longer orthogonal, and the least square approximation for the function can not be directly performed in an orthogonal sense.

We will see that this problem is solved by using the generalized Green's theorem and the construction of the adjoint operator. Examples and related properties will be shown.

2. Statement of the problem. Let $f(\overrightarrow{r})$ be a function of n variables, $\overrightarrow{r} = (x_1, ...x_n)$, and let $Q_N(\overrightarrow{r})$ be an approximating polynomial given by

$$Q_N(\overrightarrow{r}) = \sum_{k=1}^{N} a_k P_k(\overrightarrow{r})$$

where $\{P_k(\overrightarrow{r})\}$ represents a set of orthogonal polynomials in a n-dimensional volume Ω bounded by a surface C under the following inner product:

$$\int_\Omega P_k(\overrightarrow{r}) P_{k'}(\overrightarrow{r}) d\omega = \delta_{k,k'}$$

*Received December 1, 2004. Accepted for publication May 15, 2005. Recommended by J. Arvesú.
†Departamento de Física Aplicada, Facultad de Física, Universidad de Santiago de Compostela, 15782 Spain (fagoeva@usc.es).

ETNA
Kent State University
etna@mcs.kent.edu

The fact that the polynomials $P_k(\overrightarrow{r})$ are orthogonal, allow to obtain the expansion coefficients, a_k, that minimize the following norm of the difference between $f(\overrightarrow{r})$ and $Q_N(\overrightarrow{r})$

$$\|f(\overrightarrow{r}) - Q_N(\overrightarrow{r})\| = \int_\Omega (f(\overrightarrow{r}) - Q_N(\overrightarrow{r}))^2 d\omega$$

by simply performing the inner product of $f(\overrightarrow{r})$with the corresponding polynomial $P_k(\overrightarrow{r})$:

$$a_k = \int_\Omega P_k(\overrightarrow{r}) f(\overrightarrow{r}) d\omega \tag{2.1}$$

On the other hand, if L represents a linear operator (or a set of linear operators) the expansion coefficients, a_k, of the approximating function can not be obtained generally in the orthogonal sense as defined in (2.1)[1] since $\{L(P_k(\overrightarrow{r}))\}$ is no longer a set of orthogonal functions.

To solve the problem in the orthogonal sense explained above, we will make use of the generalization of the Green's Theorem [2] and the adjoint operator of L, $\widetilde{L}$. Thus if f and g are functions of $\overrightarrow{r}$ the Green's Theorem establishes that:

$$gL(f) - f\widetilde{L}(g) = \nabla \cdot \overrightarrow{V}(f, g, \partial f, \partial g, ...) \tag{2.2}$$

where $\overrightarrow{V}$ is a generalized vector depending not only on f and g but also on their partial derivatives for differential linear operators and it can be set to zero for integral operators[2].

Integrating both sides of (2.2) over Ω and making use of the divergence theorem[3] in the right side, we get

$$\int_\Omega f\widetilde{L}(g) d\omega = \int_\Omega gL(f) d\omega - \oint_C \overrightarrow{V}(f, g, \partial f, \partial g, ...) d\overrightarrow{s}$$

$d\overrightarrow{s}$ being the outward drawn vector element of area. From this expression and (2.1) it can be deduced that if there exists a set of functions $\{g_k(\overrightarrow{r})\}$such that $\widetilde{L}(g_k(\overrightarrow{r})) = P_k(\overrightarrow{r})$ then the expansion coefficient a_k can be evaluated by inner product of $L(f(\overrightarrow{r}))$ and $g_k(\overrightarrow{r})$ plus a surface integral containing boundary conditions for some cases of differential operators:

$$a_k = \int_\Omega g_k(\overrightarrow{r}) L(f(\overrightarrow{r})) d\omega - \oint_C \overrightarrow{V}(f, g_k, \partial f, \partial g_k, ...) d\overrightarrow{s}$$

3. Examples. In what follows the problem will be illustrated with two different kind of linear operators: A differential one, the gradient of a function, and an integral one, the Radon transform.

3.1. The gradient operator. If the linear operator represents the gradient of the function

$$\bigtriangledown f(\overrightarrow{r}) = \left(\frac{\partial f(\overrightarrow{r})}{\partial x_1}, ..., \frac{\partial f(\overrightarrow{r})}{\partial x_n} \right)$$

[1] Unless we wanted the approximating function to be defined in terms of the eigenfunctions of the linear operator.

ETNA
Kent State University
etna@mcs.kent.edu

the above result allow us to deduce that if there exists a set of vector functions $\{\overrightarrow{g}_k(\overrightarrow{r})\}$ such that

$$\begin{array}{c} \nabla \cdot \overrightarrow{g_k}(\overrightarrow{r}) = -P_k(\overrightarrow{r}) \text{ in } \overrightarrow{r} \in \Omega \\ \overrightarrow{g_k}(\overrightarrow{r}) \cdot \overrightarrow{n} = 0 \text{ in } \overrightarrow{r} \in C \end{array} \tag{3.1}$$

being $\overrightarrow{n}$ an unitary outward drawn vector to C, then the expansion coefficients can be evaluated independently, in the orthogonal sense, as

$$a_k = \int_\Omega \overrightarrow{g_k}(\overrightarrow{r}) \cdot \nabla f(\overrightarrow{r}) d\omega$$

Examples of vector functions orthogonal to the gradient of Zernike polynomials are obtained in a different mathematical context in references [6] and [7].

3.1.1. Properties of vector functions. In general, the set of partial differential equations given in (3.1) has not an unique solution for the set of vector function $\overrightarrow{g_k}(\overrightarrow{r})$. The choice of the set of functions, is, in principle, irrelevant, and therefore "the easier to solve the better" is a criterium to find the vector functions. Nevertheless if $\nabla f(\overrightarrow{r})$represents a physical magnitude the choice of the vector functions can improve the results. For instance let us see what happen if $\nabla f(\overrightarrow{r})$ is affected by uncorrelated noise vector,$\overrightarrow{v}$, with zero mean, and covariance $\sigma^2 \delta(\overrightarrow{r} - \overrightarrow{r}')$

Let us denote by $\widehat{a_k}$ the estimate of the coefficients when evaluated by

$$\widehat{a_k} = \int_\Omega \overrightarrow{g_k}(\overrightarrow{r}) \cdot (\nabla f(\overrightarrow{r}) + \overrightarrow{v}) d\omega$$

Then by taking the ensemble-averaged mean square error, the variance associated to the estimated coefficient is given by

$$\sigma_k^2 = \int_\Omega (\widehat{a_k} - a_k)^2 d\omega = \sigma^2 \int_\Omega \left[\overrightarrow{g_k}(\overrightarrow{r})\right]^2 d\omega \tag{3.2}$$

Thus, those vector functions minimizing (3.2) yield to a better estimate of $f(\overrightarrow{r})$. This can be get by demanding that $\overrightarrow{g_k}(\overrightarrow{r})$ satisfy not only (3.1) but also to be irrotational [8], i.e., there exists an scalar function $\xi_k(\overrightarrow{r})$ such that

$$\overrightarrow{g_k}(\overrightarrow{r}) = \nabla \xi_k(\overrightarrow{r})$$

Thus the vector functions can be found by solving the following Poisson equation with Neumann boundary condition

$$\begin{array}{c} \nabla^2 \xi_k(\overrightarrow{r}) = -P_k(\overrightarrow{r}) \text{ in } \overrightarrow{r} \in \Omega \\ \nabla \xi_k(\overrightarrow{r}) \cdot \overrightarrow{n} = 0 \text{ in } \overrightarrow{r} \in C \end{array} \tag{3.3}$$

As an example, the Zernike Polynomials formula as well as the related $\xi_k(\overrightarrow{r})$ functions obeying (3.3) is presented

ETNA
Kent State University
etna@mcs.kent.edu

$$Z_{n,l}(r,\theta) = c_{n,m} \sum_{j=0}^{\frac{n+|m|}{2}} \frac{(-1)^j (n-j)! r^{n-2j}}{j!(\frac{n+m}{2}-j)!(\frac{n-m}{2}-j)!} e^{im\theta}$$

$$\xi_{n,l}(r,\theta) = c_{n,m} \sum_{j=0}^{\frac{n+|m|}{2}} \frac{(-1)^j (n-j)!(r^{n-2j+2} - (1-\delta_{0,m})\frac{n-2j+2}{m} r^m)}{j!(\frac{n+m}{2}-j)!(\frac{n-m}{2}-j)!} e^{im\theta}$$

(the usual complex and two index notation has been used, $c_{n,m}$ are normalization constants [1] and $\delta_{0,m} = \begin{cases} 0 \text{ if m} \neq 0 \\ 1 \text{ if m} = 0 \end{cases}$)

It must be stressed the fact that depending on the set of chosen functions, $\{P_k(\overrightarrow{r})\}$, and the characteristics of the integration domain Ω, it is not always possible to obtain minimum norm solution for all the set vector functions as well as polynomic solutions as for the previous case. For instance, for orthogonal polynomials in an annulus, as C represents inner and outer circumferences, then $\nabla\xi_k(\overrightarrow{r}) \cdot \overrightarrow{n} = 0$ in $\overrightarrow{r} \in C$ represents two boundary conditions and it is not possible to find polynomic solutions for all the basis elements since terms in the shape $\log r$, $r^{-m}\sin(m\theta)$, $r^{-m}\cos(m\theta)$ are needed to fullfill both conditions. For a square or rectangular domain and a orthogonal basis made by Cartesian product of orthogonal polynomials in one dimension it is not possible to get minimum norm solution for all the basis elements due to the need of four boundary conditions.

3.2. The Radon transform in two variables. It is well-known that Radon transform can be considered as linear operator which in polar coordinates can be written as

$$Rf(p,\theta) = \int_0^{2\pi}\int_0^a f(r,\phi)K(r,\phi;p,\theta) r dr d\phi$$

being $a \in (0,\infty)$ and

$$K(r,\phi;p,\theta) = \delta(p - r\cos\phi\cos\theta - r\sin\phi\sin\theta)$$

thus representing the integral of the function $f(r,\phi)$ over the line defined in normal form by p and θ [4]

$$p - x\cos\theta - y\sin\theta = 0$$

For integral operators, $\overrightarrow{V}$ can be set to zero [2] and thus, with some algebra, it can be easily shown that the adjoint operator $\widetilde{R}g(p,\theta)$ is given by

$$\widetilde{R}f(p,\theta) = \int_0^{2\pi}\int_0^a f(r,\phi)\widetilde{K}(r,\phi;p,\theta) r dr d\phi$$

with

$$\widetilde{K}(r,\phi;p,\theta) = \frac{p}{r}K(p,\theta;r,\phi) = \frac{p}{r}\delta(r - p\cos\phi\cos\theta - p\sin\phi\sin\theta)$$

representing the integral over the circumference that is tangent to the line defined by p and θ and passes through the origin of coordinates (the tangency point and the origin define the extremes of a diameter).

In this way if there exist a set of functions $h_k(\overrightarrow{r})$ such that

$$\widetilde{R}h_k(p,\theta) = pP_k(p,\theta)$$

then

$$a_k = \int_0^{2\pi}\int_0^{a} h_k(p,\theta)Rf(p,\theta)dpd\theta$$

For $a = 1$ and choosing as orthogonal basis the Zernike polynomials, the $h_k(\overrightarrow{r})$ functions are related to Chebyshev polynomials of second kind[4][2]

As an example the first six orthogonal polynomials of Appell and Kampé de Fériet[5] for the unit circle and the corresponding $h_k(\overrightarrow{r})$ functions are shown

$U_k(p,\theta)$	$h_k(p,\theta)$
1	$\frac{1}{\pi}$
$p\cos\theta$	$\frac{2}{\pi}p\cos\theta$
$p\sin\theta$	$\frac{2}{\pi}p\sin\theta$
$3p^2\cos^2\theta + p^2\sin^2\theta - 1$	$\frac{8}{\pi}p^2\cos^2\theta - \frac{1}{\pi}$
$p^2\sin\theta\cos\theta$	$\frac{4}{\pi}p\sin\theta\cos\theta$
$3p^2\sin^2\theta + \cos^2\theta - 1$	$\frac{8}{\pi}p^2\sin^2\theta - \frac{1}{\pi}$

4. Conclusions. By the use of the generalized Green's theorem it is possible to find set of functions that are orthogonal to the action of linear operators upon a set of orthogonal polynomials. This is equivalent to solve linear equations, both differential equations and integral equations, in an orthogonal least squares sense without solving an eigenfunctions problem.

For the gradient operator the set of functions orthogonal to the gradient of the elements of the chosen basis is not unique, and it has been shown that in some cases it is possible to find only one set obeying additional conditions, that minimize the norm of the difference between the function and its approximating polynomial if the gradient of the function is affected by gaussian noise of zero mean and spatially uncorrelated.

For linear operators, as Radon transform, since no boundary conditions are needed, the set is unique and depends only on the chosen basis and the chosen integration domain

The method can be extended to inner products with weights different than unity as well as to orthogonal functions in general.

Acknowledgements. This work was supported by Ministerio de Educción y Ciencia, Grant No. AYA 2004-07773-C02-02.

[2]The orthogonality between Radon transform of Zernike polynomials and Chebyshev polynomials need integration in the radial coordinate between -1 and 1

REFERENCES

[1] M. BORN AND E. WOLF, *Principles of Optics*, Pergamon Press, New York, 1986.
[2] P.M. MORSE AND H. FESHBACH, *Methods of Theoretical Physics*, McGraw-Hill Book Co., Inc., New York-Toronto-London, 1953, pp. 999–1978.
[3] H.F. WEIMBERGER, *Partial Differential Equations*, Blaisdell Publishing, New York, 1965.
[4] R.B. MARR, *On the reconstruction of a function on a circular domain from a sampling of its line integrals*, J. Math. Anal. Appl., 45 (1974), pp. 357–374.
[5] P. APPELL AND J. KAMPÉ DE FÉRIET, *Fontions Hypergéomètriques et hyperspheriques-Polynomes d'Hermite*, Gauthier-Villars, Paris, 1926.
[6] V.P. AKSENOV AND YU. N. ISAEV, *Analytical representation of the phase and its mode components reconstructed according to the wave-front slopes*, Opt. Lett., 17 (1992), 1180.
[7] E. ACOSTA, S. BARÁ, M.A. RAMA AND S. RÍOS, *Determination of phase mode components in terms of local wave-front slopes: an analytical approach*, Opt. Lett., 20 (1995), 1083.
[8] C.J. SOLOMON, S. RÍOS, E. ACOSTA AND S. BARÁ, *Modal wavefront projectors of minimum error norm*, Opt. Commun., 155 (1998), 251.