8. The analogy between matrices and operators

In the typical problem of linear algebra, we solve

(1) $\mathbf{Lu} = \mathbf{f}$

for $\mathbf{u}$, where $\mathbf{L}$ is an $n \times n$ matrix, and $\mathbf{u}$ and $\mathbf{f}$ are $(n \times 1)$ column vectors. This problem has a unique solution if $\det \mathbf{L} \neq 0$. One method—usually not the best method—of solving (1) is to first find the inverse $\mathbf{L}^{-1}$ of $\mathbf{L}$, that is, to solve

(2) $\mathbf{L}\left(\mathbf{L}^{-1}\right) = \mathbf{I}$

for $\mathbf{L}^{-1}$ and then compute $\mathbf{u}=\mathbf{L}^{-1}\mathbf{f}$. A better method—though not always possible—is to transform the variables into a new system in which $\mathbf{L}$ takes a simple form. The ultimate simplification is one in which the transformed $\mathbf{L}$ is purely diagonal.

Our interest is differential equations. Hence we wish to solve

(3) $Lu(\mathbf{x}) = f(\mathbf{x})$ with the appropriate boundary conditions

where L is a linear partial differential operator, and $u(\mathbf{x})$ and $f(\mathbf{x})$ are functions. The linear algebra problem (1) and the system (3) are analogous. The analogy is between:

$\{u_1, u_2, \ldots, u_n\}$, the n components of $\mathbf{u}$; and $\{u(\mathbf{x}),\ \text{all } \mathbf{x}\}$, an infinite set

$\{f_1, f_2, \ldots, f_n\}$, and $\{f(\mathbf{x}),\ \text{all } \mathbf{x}\}$

$\mathbf{L}$, an $n \times n$ matrix; and L, an $\infty \times \infty$ operator.

In this section we explore the analogy between (1) and (3). The main benefit of this is a better understanding of the relationship between the various solution strategies for (3).

In the "transformation approach" to simplifying (1), one lets

(4) $\mathbf{u} = \mathbf{T}\hat{\mathbf{u}}$, i.e. $\hat{\mathbf{u}} = \mathbf{T}^{-1}\mathbf{u}$

where $\mathbf{T}$ is an $n \times n$ matrix and $\hat{\mathbf{u}}$ are the new variables. (We require $\det \mathbf{T} \neq 0$ for an invertible transformation.) Then (1) takes the form $\mathbf{LT}\hat{\mathbf{u}} = \mathbf{T}\hat{\mathbf{f}}$, which implies

(5) $\hat{\mathbf{L}}\hat{\mathbf{u}} = \hat{\mathbf{f}}$ where $\hat{\mathbf{L}} = \mathbf{T}^{-1}\mathbf{LT}$ is the transformation of $\mathbf{L}$.

The strategy is to find a $\mathbf{T}$ that makes $\hat{\mathbf{L}}$ simpler than $\mathbf{L}$. Special properties of $\mathbf{L}$ may allow a spectacular simplification.

We rewrite (4) as

(6) $$u_i = \sum_{j=1}^{n} T_{ij}\ \hat{u}_j,$$ which is analogous to

(7) $$u(\mathbf{x}) = \iiint d\mathbf{x}'\ T(\mathbf{x},\mathbf{x}')\ \hat{u}(\mathbf{x}').$$

The arrows point to the corresponding things. For example,

(8) $$u(\mathbf{x}) = \iiint d\mathbf{k}\ e^{i\mathbf{k}\cdot\mathbf{x}}\ \hat{u}(\mathbf{k})$$

fits the form of (7) with $T(\mathbf{x},\mathbf{k}) = e^{i\mathbf{k}\cdot\mathbf{x}}$. The analog of $\hat{\mathbf{u}} = \mathbf{T}^{-1}\mathbf{u}$ is

(9) $$\hat{u}(\mathbf{k}) = \iiint d\mathbf{x}\ \frac{1}{(2\pi)^3} e^{-i\mathbf{k}\cdot\mathbf{x}}\ u(\mathbf{x}).$$

Thus

(10) $$T^{-1}(\mathbf{k},\mathbf{x}) = \frac{1}{(2\pi)^3} e^{-i\mathbf{k}\cdot\mathbf{x}}.$$

The (Fourier) transformation from $u(\mathbf{x})$ to $\hat{u}(\mathbf{k})$ is useful for all equations with constant coefficients in unbounded domains. For example, it transforms

$\nabla^2 u = Q$ into $-\mathbf{k}\cdot\mathbf{k}\,\hat{u} = \hat{Q}$, a completely diagonal form.

Next we ask: What is the analogue of matrix inversion? Clearly the analogue of

(11) $$u_i = \sum_j L_{ij}^{-1} f_j$$

is

(12) $$u(\mathbf{x}) = \iiint d\mathbf{x}'\ G(\mathbf{x},\mathbf{x}')\ f(\mathbf{x}').$$

Thus the Green's function is analogous to the inverse matrix $\mathbf{L}^{-1}$. Applying the operator L to (12) we obtain

(13) $$Lu(\mathbf{x}) = \iiint d\mathbf{x}'\ LG(\mathbf{x},\mathbf{x}')\ f(\mathbf{x}')$$

which agrees with (3) provided that

(14) $$LG(\mathbf{x},\mathbf{x}') = \delta(\mathbf{x}-\mathbf{x}').$$

(Note that L operates on $\mathbf{x}$ and *not* $\mathbf{x}'$.) Thus (14) appears to be the analogue of

(15) $$\mathbf{L}\ \mathbf{L}^{-1} = \mathbf{I}.$$

This is not quite true! While we know that (15) has a unique solution for $\mathbf{L}^{-1}$ provided that $\det \mathbf{L} \neq 0$, (14) may be solved uniquely for G only if it is accompanied by the approporiate boundary conditions. In other words, the operator L is invertible only if it is defined in such a way as to incorporate the boundary conditions.

To clarify this we first note that the analogy between $\mathbf{L}$ and L can be turned into an equivalence if we agree to replace L by a finite-difference approximation. To keep things as simple as possible, suppose that $L = \partial^2 / \partial x^2$ so that our equation is the one-dimensional Poisson equation,

$$u_{xx} = f. \tag{16}$$

The finite-difference version is

$$\frac{u_{i+1} + u_{i-1} - 2u_i}{\Delta^2} = f_i \tag{17}$$

where Δ is the spacing between gridpoints. If we let $\mathbf{u}$ be the column vector of gridded values, then, on the infinite domain, (16) takes the form

$$\begin{bmatrix} \ddots & & & & \\ & & & & \\ & \frac{1}{\Delta^2} & \frac{-2}{\Delta^2} & \frac{1}{\Delta^2} & \\ & & & & \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ u_{i-1} \\ u_i \\ u_{i+1} \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ f_{i-1} \\ f_i \\ f_{i+1} \\ \vdots \end{bmatrix} \tag{18}$$

We see that $\mathbf{L}$ is an $\infty \times \infty$ tridiagonal matrix. We also see that $\mathbf{L}$ must be singular, because we can add a constant to any solution of (18) and obtain another solution. (This corresponds to one null vector of $\mathbf{L}$. What is the other null vector?)

If we make the domain finite and impose Dirichlet boundary conditions at the endpoints, then we obtain the *nonsingular* problem

$$\begin{bmatrix} 1 & & & & & & \\ 1/\Delta^2 & -2/\Delta^2 & 1/\Delta^2 & & & & \\ & & \ddots & \ddots & & & \\ & & & 1/\Delta^2 & -2/\Delta^2 & 1/\Delta^2 \\ & & & & & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{n-1} \\ u_n \end{bmatrix} = \begin{bmatrix} 0 \\ f_2 \\ \vdots \\ f_{n-1} \\ 0 \end{bmatrix} \tag{19}$$

The matrix $\mathbf{L}$ in (19) is the analogue of L *plus* the boundary conditions u=0. Let $\mathbf{L}^{-1} \equiv \mathbf{G}$ be the inverse of the matrix in (19). To find $\mathbf{G}$ we must solve

$$\mathbf{LG} = \mathbf{I}. \tag{20}$$

We may solve for $\mathbf{G}$ "one column at a time" by writing

(21) $$\begin{bmatrix} 1 & & & & & \\ 1/\Delta^2 & -2/\Delta^2 & 1/\Delta^2 & & & \\ & & \ddots & \ddots & & \\ & & & 1/\Delta^2 & -2/\Delta^2 & 1/\Delta^2 \\ & & & & & 1 \end{bmatrix} \begin{bmatrix} G_{1j} \\ G_{2j} \\ \vdots \\ G_{n-1,j} \\ G_{nj} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

where the 1 on the right-hand side appears in the j-th row. We recognize (21) as the finite-difference analogue of

(22) $$LG(x,x') = \delta(x - x')$$

and the boundary conditions

(23) $$G(0,x') = G((n-1)\Delta, x') = 0 .$$

This proves that the inversion operation includes the imposition of the same boundary conditions on G that were applied to u. Of course solving (21) by (say) Gaussian elimination requires n times as much work as solving (19) directly for **u**. Thus Green's functions are impractical unless the symmetry of the problem makes them easy to find.

Everything we have said so far applies to general L and **L**. Now we consider the special case of symmetric **L**. If **L** is symmetric, then we know from section 5 that its eigenvalues are real, and the corresponding eigenvectors are (or can be made) orthogonal. This means that the eigenvectors span the whole n-dimensional space. Thus any column vector, including **u** and **f**, can be represented as a weighted sum of the eigenvectors. *The eigenvectors form a complete set*. Furthermore, if **L** is symmetric then so is $\mathbf{L}^{-1}$. Finally, if **L** is symmetric, then

(24) $$\sum_j L_{ij}\, u_j = f_i$$

may be written

(25) $$\frac{\partial}{\partial u_k}\left(\sum_i \sum_j \frac{1}{2} u_i\, L_{ij}\, u_j - \sum_i f_i\, u_i \right) = 0 \qquad \text{for } k\text{=1 to } n.$$

In other words, the problem **Lu=f** is equivalent to the *variational principle*

(26) $$\delta\left(\frac{1}{2}\mathbf{u}^T\mathbf{L}\mathbf{u} - \mathbf{u}^T\mathbf{f} \right) = 0$$

where δ stands for arbitrary independent variations in the components of the column vector **u**. These facts provide the motivation to seek the analogous properties for operators that are the analogues of symmetric matrices. Such operators are called *self-adjoint*. But how exactly do we define self-adjointness?

One approach is this: If the matrix corresponding to the finite-difference form of the differential equation is symmetric, then the operator is said to be self-adjoint. For example, we note that the matrix in (19) is *not* symmetric (but only because of its first and last rows). Thus the problem

$$(27)\qquad \begin{aligned} &u_{xx} = f, \qquad 0 < x < 1 \\ &u(0) = u(1) = 0 \end{aligned}$$

is not self-adjoint in the sense defined above. Suppose however that we change the boundary conditions to $u_x(0) = u_x(1) = 0$. Then the matrix in (19) becomes

$$(28)\qquad \begin{bmatrix} -1/\Delta^2 & 1/\Delta^2 & & & & \\ 1/\Delta^2 & -2/\Delta^2 & 1/\Delta^2 & & & \\ & & \ddots & \ddots & & \\ & & & 1/\Delta^2 & -2/\Delta^2 & 1/\Delta^2 \\ & & & & 1/\Delta^2 & -1/\Delta^2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{n-1} \\ u_n \end{bmatrix} = \begin{bmatrix} 0 \\ f_2 \\ \vdots \\ f_{n-1} \\ 0 \end{bmatrix}$$

which *is* symmetric. Thus the problem

$$(29)\qquad \begin{aligned} &u_{xx} = f, \qquad 0 < x < 1 \\ &u_x(0) = u_x(1) = 0 \end{aligned}$$

is self-adjoint. But wait! If we simply use the fact that $u_1 = u_n = 0$ to rewrite (19) as an $(n-2) \times (n-2)$ matrix equation,

$$(30)\qquad \begin{bmatrix} -1/\Delta^2 & 1/\Delta^2 & & & & \\ 1/\Delta^2 & -2/\Delta^2 & 1/\Delta^2 & & & \\ & & \ddots & \ddots & & \\ & & & 1/\Delta^2 & -2/\Delta^2 & 1/\Delta^2 \\ & & & & 1/\Delta^2 & -1/\Delta^2 \end{bmatrix} \begin{bmatrix} u_2 \\ u_3 \\ \vdots \\ u_{n-2} \\ u_{n-1} \end{bmatrix} = \begin{bmatrix} f_2 \\ f_3 \\ \vdots \\ f_{n-2} \\ f_{n-1} \end{bmatrix}$$

then the lower-dimensional matrix is symmetric. So is the problem (27) self-adjoint or not? The answer clearly depends on the size of the function space. If the function space includes only those $u(x)$ that vanish at the endpoints, then the problem is self-adjoint.

This agrees with something we discovered in Section 7. If (27) is self-adjoint it should have a variational principle. The variational principle is in fact

$$(31)\qquad \delta \int_0^1 dx \left(\frac{1}{2} u_x^{\,2} + f\,u \right) = 0.$$

But the variational principle (31) only works if we constrain $u = \delta u = 0$ at the endpoints. If, on the other hand, we place no restrictions on u and δu at the endpoints, then the variational principle gives us the problem (29) with its boundary conditions of $u_x = 0$. The matrix in (28) corresponding to this problem is symmetric. Thus the problem (29) is self-adjoint on the *unrestricted* function space.

If self-adjointness really matters, then it is disturbing that it should depend on the manner in which we choose to incorporate the boundary conditions. It seems better to redefine self-adjointness as a property of the operator alone, and to leave the boundary conditions as a separate issue. Moreover, it seems a bad idea to define properties of the operator based on the form of the corresponding finite-difference equations. Not only are these non-unique, but we note that multiplying any row of (19) or (28) by a constant, or interchanging any 2 rows, would destroy the symmetry property without changing the problem at all. Finally, writing out the finite-differences in more than 1 space dimension becomes very tedious. Considering all this, we are on very thin ice!

We therefore proceed by manipulations that do not refer to matrices or finite-differences at all, but simply use the matrix theory as a guide. That is, we pursue the analogy withou attempting to show an equivalence.

Definition. The operator L is self-adjoint if

$$\int dx\ v\,Lu = \int dx\ u\,Lv \ + \ \text{boundary terms} \tag{32}$$

for any 2 functions $u(x)$ and $v(x)$.

This is clearly the analogue of the statement $\mathbf{u}^T\mathbf{L}\mathbf{v} = \mathbf{v}^T\mathbf{L}\mathbf{u}$ for any 2 column vectors **u** and **v**, which in turn implies that $\mathbf{L} = \mathbf{L}^T$. However the definition (32) does not require us to determine the "components" of L by forming finite differences. By this definition $L = \partial^2/\partial x^2$ is self-adjoint. More generally, $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ is self-adjoint within any closed domain. To prove this we note that

$$\begin{aligned} u\nabla^2 v &= \nabla\cdot(u\nabla v) - \nabla u\cdot\nabla v \\ v\nabla^2 u &= \nabla\cdot(v\nabla u) - \nabla u\cdot\nabla v \end{aligned} \tag{33}$$

Subtracting and integrating we obtain

$$\iint dx\,dy\,\left\{u\nabla^2 v - v\nabla^2 u\right\} = \iint dx\,dy\ \nabla\cdot(u\nabla v - v\nabla u). \tag{34}$$

By the divergence theorem this becomes

$$\iint dx\,dy\,\left\{u\nabla^2 v - v\nabla^2 u\right\} = \oint ds\,\left(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n}\right), \tag{35}$$

which fits the form of (32).

Consider the problem

$$\nabla^2 u = f(x,y) \tag{36}$$

in an arbitrary domain with mixed boundary condition

$$\alpha u + \beta\frac{\partial u}{\partial n} = 0 \tag{37}$$

where α and β are prescribed functions. (This includes Neumann and Dirichlet boundary conditions as the special cases α=0 and β=0.)

Now we state and check properties of (36-37) which are expected from the analogy with the matrix problem:

• Since ∇^2 is self-adjoint, we expect its eigenfunctions to be orthogonal. To see that they are, let u_1 and u_2 be any 2 eigenfunctions. Then

$$\text{(38)} \qquad \begin{aligned} \nabla^2 u_1 &= \lambda_1 u_1 \\ \nabla^2 u_2 &= \lambda_2 u_2 \end{aligned} \qquad \text{with boundary conditions} \qquad \begin{aligned} \alpha u_1 + \beta \frac{\partial u_1}{\partial n} &= 0 \\ \alpha u_2 + \beta \frac{\partial u_2}{\partial n} &= 0 \end{aligned}$$

where λ_1 and λ_2 are the eigenvalues. Multiplying (38a) by u_2, (38b) by u_1, subtracting, integrating and using the property (35) we obtain

$$\text{(39)} \qquad \oint ds \left(u_2 \frac{\partial u_1}{\partial n} - u_1 \frac{\partial u_2}{\partial n} \right) = \left(\lambda_1 - \lambda_2 \right) \iint u_1 u_2 \; dx\, dy .$$

the left-hand side vanishes because of the boundary conditions (37). Thus if $\lambda_1 \neq \lambda_2$, the eigenfunctions are indeed orthogonal. (Note how this proof mirrors the corresponding proof for matrices in Section 5.)

• Since ∇^2 is self-adjoint [analogous statement: $\mathbf{L}$ is symmetric], we expect its Green's function to be symmetric in its arguments [analogous statement: $\mathbf{L}^{-1}$ is symmetric]. The Green's function corresponding to (36-37) obeys

$$\text{(40)} \qquad \nabla^2 G(\mathbf{x},\mathbf{x}_0) = \delta(\mathbf{x} - \mathbf{x}_0) \quad \text{and the boundary conditions} \quad \alpha G + \beta \frac{\partial G}{\partial n} = 0 .$$

From (40) it follows that

$$\text{(41)} \qquad u(\mathbf{x}) = \iint d\mathbf{x}_0 \; G(\mathbf{x},\mathbf{x}_0) f(\mathbf{x}_0) .$$

To show that G is symmetric we multiply (40) by $u(\mathbf{x})$, and (36) by $G(\mathbf{x},\mathbf{x}_0)$, and proceed as toward (39) to obtain

$$\text{(42)} \qquad \oint ds \left(u \frac{\partial G}{\partial n} - G \frac{\partial u}{\partial n} \right) = \iint \; d\mathbf{x} \; \left\{ u(\mathbf{x}) \delta(\mathbf{x} - \mathbf{x}_0) - G(\mathbf{x},\mathbf{x}_0) f(\mathbf{x}) \right\}$$

The left-hand side vanishes because of the boundary conditions on G and u. Thus

$$\text{(43)} \qquad u(\mathbf{x}_0) = \iint d\mathbf{x} \; G(\mathbf{x},\mathbf{x}_0) f(\mathbf{x}) .$$

By interchanging $\mathbf{x}_0$ and $\mathbf{x}$ in (43) and comparing the result with (41), we see that

(44) $$G(\mathbf{x},\mathbf{x}_0) = G(\mathbf{x}_0,\mathbf{x})$$

as expected.

• Since ∇^2 is self-adjoint, we expect the problem (36-37) to correspond to a variational principle. From Section 7 we know that it does.

• Finally since ∇^2 is self-adjoint, we expect that its eigenfunctions form a complete set of functions. (This is analogous to the statement that the eigenvectors of symmetric **L** span the whole *n*-dimensional space.) Unfortunately, completeness is a property that does not *automatically* accompany self-adjointness; it depends on the particular case. However, we can prove that the eigenfunctions of ∇^2 are complete using tools that arise from the self-adjointness.

Sketch of the proof.

For simplicity we assume Dirichlet boundary conditions, but the proof can be extended to Neumann or mixed boundary conditions. By definition the eigenfunctions obey

(45) $$\nabla^2 u + \lambda u = 0 \qquad \text{with boundary condition } u=0.$$

For example, if the domain is $0 < x, y < \pi$, then the eigenfunctions are $\theta_{nm}(x,y) = \sin nx \ \sin my$ with eigenvalues $\lambda_{nm} = n^2 + m^2$.

In general, the normalized eigenfunction u_1 with the smallest eigenvalue λ_1 is the function that minimizes

(46) $$\iint \nabla u \cdot \nabla u$$

subject to the constraints

(47) $$\iint u^2 = 1 \qquad \text{and } u=0 \text{ at the boundary.}$$

That is, of all the functions satisfying (47), u_1 is the one that minimizes (46). Similarly, the eigenfunction u_2 with the next-smallest eigenvalue is the function that minimizes (46) subject to

(48) $$\iint u^2 = 1, \qquad \iint u u_1 = 0, \qquad \text{and } u=0 \text{ at the boundary.}$$

Proceeding in the obvious way, the *n*-th eigenfunction minimizes (46) subject to

(49) $\iint u^2 = 1, \qquad \iint u u_1 = \cdots \iint u u_{n-1}, \qquad$ and u=0 at the boundary.

Within the set of functions satisfying (49) u_n gives (46) its minimum value.

To see that this procedure is equivalent to (45), consider the "amplitude-insensitive" functional

(50) $$R[u] = \frac{\iint \nabla u \cdot \nabla u}{\iint u^2}$$

We shall show that solutions of (45) are stationary points of (50). Using u=δu=0 at the boundary,

(51) $$\delta R = 2\frac{\iint \nabla u \cdot \nabla \delta u}{\iint u^2} - \frac{\iint \nabla u \cdot \nabla u}{\left(\iint u^2\right)^2} 2\iint u\,\delta u = 0$$

implies

(52) $$\iint \left(-\nabla^2 u - \lambda\, u\right)\delta u = 0$$

where

(53) $$\lambda = \frac{\iint \nabla u \cdot \nabla u}{\iint u^2}.$$

Thus the variational principle δR=0 yields (45); and, moreover, the value of R at the solution of (45) is the eigenvalue λ.

Returning to the subject of completeness, we let $f(x,y)$ be an arbitrary function satisfying the boundary condition f=0. If the eigenfunctions are complete, then

(54) $$f(x,y) = \sum_{i=1}^{\infty} c_i u_i(x,y) \qquad \text{where} \qquad c_i = \iint f\, u_i .$$

To test this we define the partial sum

(55) $$f_n(x,y) \equiv \sum_{i=1}^{n} c_i\, u_i(x,y)$$

and the remainder $r_n(x,y) \equiv f - f_n$. Our task is to show that $r_n \to 0$ (in some sense) as $n \to \infty$. But r_n belongs to the class (49) except for the normalization requirement. Thus

(56) $$R\left[r_n\right] \geq \lambda_{n+1} \qquad \Leftrightarrow \qquad \frac{\iint \nabla r_n \cdot \nabla r_n}{\iint r_n^{\;2}} \geq \lambda_{n+1}$$

The proof proceeds by showing that $\iint \nabla r_n \cdot \nabla r_n$ is bounded as $n \to \infty$. Then since $\lambda_{n+1} \to \infty$, we must have $\iint r_n^{\;2} \to 0$, i.e.

(57) $$\iint \left(f - \sum_{i=1}^{n} c_i\, u_i \right)^2 \to 0 \qquad \text{as} \quad n \to \infty$$

(convergence in the mean square). To see that $\iint \nabla r_n \cdot \nabla r_n$ is bounded, we compute

$$\iint \nabla r_n \cdot \nabla r_n = \iint \left[\nabla f \cdot \nabla f - 2\nabla f \cdot \nabla f_n + \nabla f_n \cdot \nabla f_n \right]$$

$$= \iint \left[\nabla f \cdot \nabla f + 2 f \nabla^2 f_n - f_n \nabla^2 f_n \right]$$

(58) $$= \iint \left[\nabla f \cdot \nabla f - 2 f \sum_{i=1}^{n} \lambda_i c_i u_i + \sum_{i=1}^{n} \sum_{i=1}^{n} \; c_i u_i \lambda_j c_j u_j \right]$$

$$= \iint \nabla f \cdot \nabla f - 2 \sum_{i=1}^{n} \lambda_i c_i^2 + \sum_{i=1}^{n} \lambda_i c_i^{\;2}$$

$$= \iint \nabla f \cdot \nabla f - \sum_{i=1}^{n} \lambda_i c_i^{\;2} < \iint \nabla f \cdot \nabla f$$

Since only mean square convergence has been proved, we can even relax the condition that f=0 at the boundaries. That is, *any* $f(x,y)$ can be represented as an infinite series of the eigenfunctions in a mean-square sense.

Reference. The book by Lanczos.