Useful Finite Difference Approximations

Table A–1 First-Order Finite Difference Approximations

Difference	First Order Formula	
Forward Difference for First Derivative	$\frac{d}{dx}f(x_i) = \frac{f(x_{i+1}) - f(x_i)}{\Delta x}$	(A-1)
Backward Difference for First Derivative	$\frac{d}{dx}f(x_i) = \frac{f(x_i) - f(x_{i-1})}{\Delta x}$	(A-2)
Forward Difference for Second Derivative	$\frac{d^2}{dx^2}f(x_i) = \frac{f(x_i) - 2f(x_{i+1}) + f(x_{i+2})}{\Delta x^2}$	(A-3)
Backward Difference for Second Derivative	$\frac{d^2}{dx^2}f(x_i) = \frac{f(x_i) - 2f(x_{i-1}) + f(x_{i-2})}{\Delta x^2}$	(A-4)

Table A–2 Second-Order Finite Difference Approximations

Difference	Second Order Formula	
Forward Difference for First Derivative	$\frac{d}{dx}f(x_i) = \frac{-3f(x_i) + 4f(x_{i+1}) - f(x_{i+2})}{2\Delta x}$	(A-5)
Central Difference for First Derivative	$\frac{d}{dx}f(x_i) = \frac{f(x_{i+1}) - f(x_{i-1})}{2\Delta x}$	(A-6)
Backward Difference for First Derivative	$\frac{d}{dx}f(x_i) = \frac{3f(x_i) - 4f(x_{i-1}) + f(x_{i-2})}{2\Delta x}$	(A-7)
Forward Difference for Second Derivative	$\frac{d^2}{dx^2}f(x_i) = \frac{2f(x_i) - 5f(x_{i+1}) + 4f(x_{i+2}) - f(x_{i+3})}{\Delta x^2}$	(A-8)
Central Difference for Second Derivative	$\frac{d^2}{dx^2}f(x_i) = \frac{f(x_{i+1}) - 2f(x_i) + f(x_{i-1})}{\Delta x^2}$	(A-9)
Backward Difference for Second Derivative	$\frac{d^2}{dx^2}f(x_i) = \frac{2f(x_i) - 5f(x_{i-1}) + 4f(x_{i-2}) - f(x_{i-3})}{\Delta x^2}$	(A-10)

Abstract

Classnotes on basic properties of finite difference methods for time dependent linear PDEs.

1 Convergence, Consistency, Stability

General form of *time dependent linear* PDE (Partial Differential Equation) is:

$$P(\partial_t, \partial_x)u = f(t,x). \tag{1.1}$$

Examples are:

$$u_t = bu_{xx},\ b > 0, \tag{1.2}$$

heat equation for real function $u = u(t,x)$ with heat conductivity $b > 0$;

$$u_t + au_x = 0, \tag{1.3}$$

wave equation for real function $u = u(t,x)$ with velocity a;

$$iu_t = bu_{xx}, \tag{1.4}$$

b nonzero real number, u a complex function, (1.4) is known as Schrödinger equation. The initial value of these equations is $u_0(x)$.

One discretizes (1.1) by approximating time and spatial derivatives by finite difference quotients with grid size h in space, grid size k in time. The resulting numerical solution is denoted by v_m^n approximating $u(mh, nk)$.

Definition 1.1 *(Convergence) A one step method is convergent if for any solution $u(t,x)$ to the PDE (1.1), and numerical solution v_m^n such that v_m^0 ocnverges to $u_0(x)$ as $mh \to x$, then v_m^n converges to $u(t,x)$ as $(nk, mh) \to (t,x)$ as h, k tend to zero.*

*Department of Mathematics and TICAM, University of Texas at Austin, Austin, TX 78712. Email: jxin@math.utexas.edu.

A commonly used norm to measure convergence for a discrete grid function w is the L^2 (energy) norm:

$$\|w\|_h = \left(h \sum_{m=-\infty}^{\infty} |w_m|^2 \right)^{1/2}. \tag{1.5}$$

Definition 1.2 *(Consistency and Order of Accuracy) Given $P_{k,h}v = f$ a finite difference method, and PDE $Pu = f$, the difference method is consistent with the PDE if for any smooth function $\varphi = \varphi(t,x)$:*

$$P\varphi - P_{h,k}\varphi \to 0,$$

as k, $h \to 0$, convergence being pointwise. If in addition,

$$P\varphi - P_{h,k}\varphi = O(k^p) + O(h^q),$$

the method is accurate of order (p,q).

This is same as substituting the exact solution into difference method and expecting error terms to tend to zero with grid refinement.

Definition 1.3 *(Stability) A one step finite difference method $P_{k,h}v_m^n = 0$ for a first order equation is stable if there are positive numbers h_0, k_0 such that for any positive time T, one can find a constant C_T so that:*

$$\|v_m^n\|_h \le C_T \|v_m^0\|_h, \tag{1.6}$$

for $0 \le nk \le T$, $0 < h \le h_0$, $0 < k \le k_0$.

A beautiful characterization of convergence is given in:

Theorem 1.1 *(Lax-Richtmyer Equivalence) A consistent finite difference method for a well-posed initial value problem of linear PDE (1.1) is convergent if and only if it is stable.*

Here well-posedness of (1.1) means that:

$$\|u(t,x)\|_2 \le C_T \|u(0,x)\|_2, \; t \in [0,T]. \tag{1.7}$$

For linear PDE, stability turns out to be easier to check than convergence, thanks to Fourier analysis.

Definition
The $L2$-norm of a grid function w, denoted by $\|w\|_{\Delta x}$, is defined as

$$\|w\|_{\Delta x} = \left(\Delta x \sum_{m=-\infty}^{\infty} |w_m|^2 \right)^{1/2}.$$

Definition
A finite difference scheme $P_{\Delta t, \Delta x} v_i^n = 0$ for a first-order equation is **stable** in a stability region Λ if there is an integer J such that for any positive time T, there is a constant C_T such that

$$\|v^n\|_{\Delta x} \leq C_T \sum_{j=0}^{J} \|v^j\|_{\Delta x}$$

for $0 \leq n\Delta t \leq T$, with $(\Delta t, \Delta x) \in \Lambda$.

Von Neumann Analysis

Proving stability directly from the definition is quite difficult, in general. Instead, it is easier to use tools from Fourier analysis to evaluate the stability of finite difference schemes. In particular, it can be shown that, for some solution to a finite difference scheme v^n, there is a simple mathematical relationship between the Fourier transforms $\hat{v}^n(\xi)$ and $\hat{v}^0(\xi)$ given by

$$\hat{v}^n(\xi) = g(\Delta x \xi, \Delta t, \Delta x)^n \hat{v}^0(\xi)$$

where $g(\Delta x \xi, \Delta t, \Delta x) = g(\theta, \Delta t, \Delta x)$ is called the *amplification factor* and $\hat{v}^n(\xi)$ is the amplitude of the frequency ξ in the solution v^n (note that the second superscript n in the above equation is a power and not an index). This quantity is so named because it represents the amount that each frequency in the solution is amplified in advancing the solution one time step. That this relationship arises from such analysis should not be surprising: the fundamental power of Fourier transforms is that they transform differentiation in the temporal domain to multiplication in the frequency domain.

Definition
A one-step finite difference scheme with constant coefficients is stable in a stability region Λ if and only if there is a constant K (independent of θ, Δt, and Δx) such that

$$|g(\theta, \Delta t, \Delta x)| \leq 1 + K\Delta t$$

with $(\Delta t, \Delta x) \in \Lambda$. If $g(\theta, \Delta t, \Delta x)$ is independent of Δt and Δx, the stability condition may be replaced with the restricted stability condition

$$|g(\theta)| \leq 1.$$

Example
Consider again the one-way wave equation and let us evaluate the stability of the forward-time forward-space scheme given in equation (1). It can be shown that all solutions of any one-step difference scheme will have the form $v_j^n = g^n e^{ij\theta}$. To analyze the stability of such a scheme, we may therefore substitute accordingly and solve for g.

$$\frac{g^{n+1}e^{ij\theta} - g^n e^{ij\theta}}{\Delta t} + \alpha \frac{g^n e^{i(j+1)\theta} - g^n e^{ij\theta}}{\Delta x} = 0$$

$$\frac{g^n e^{ij\theta}(g-1)}{\Delta t} = -\alpha \frac{g^n e^{ij\theta}(e^{i\theta} - 1)}{\Delta x}$$

$$(g-1) = -\alpha(e^{i\theta} - 1)\frac{\Delta t}{\Delta x}$$

$$g = 1 + \alpha \frac{\Delta t}{\Delta x} - \alpha \frac{\Delta t}{\Delta x} e^{i\theta}$$

$$g = 1 + \alpha\lambda - \alpha\lambda e^{i\theta}$$

where $\lambda = \Delta t/\Delta x$ and α is positive. We can then calculate

$$|g|^2 = 1 + 4\alpha\lambda(1 + \alpha\lambda)\sin^2 \frac{1}{2}\theta.$$

Since λ is constant, we may use the restricted stability condition and we see that $|g|$ is greater than 1 for $\theta \neq 0$. Therefore, this scheme is unstable. □

The Lax-Richtmyer Equivalence Theorem

The Lax-Richtmyer Equivalence Theorem is often called the *Fundamental Theorem of Numerical Analysis*, even though it is only applicable to the small subset of linear numerical methods for well-posed, linear partial differential equations. Along with Dahlquist's equivalence theorem for ordinary differential equations, the notion that the relationship

$$\text{consistency} + \text{stability} \iff \text{convergence}$$

always holds has caused a great deal of confusion in the numerical analysis of differential equations. In the case of PDEs, mathematicians are most often interested in nonlinear phenomena, for which Lax-Richtmyer does not apply. More damningly, the forward implication that consistency + stability $\implies$ convergence is trivial for linear schemes, and thus it is only the converse notion that convergence $\implies$ stability that the theorem contributes. The intuition that the theorem gives for problems that fall outside the scope of Lax-Richtmyer, however, is faulty, since consistency and stability are often insufficient for convergence, and convergence need not imply stability in general.

Numerical stability [edit]

The stability of numerical schemes is closely associated with numerical error. A finite difference scheme is stable if the errors made at one time step of the calculation do not cause the errors to increase as the computations are continued. A neutrally stable scheme is one in which errors remain constant as the computations are carried forward. If the errors decay and eventually damp out, the numerical scheme is said to be stable. If, on the contrary, the errors grow with time the numerical scheme is said to be unstable. The stability of numerical schemes can be investigated by performing von Neumann stability analysis. For time-dependent problems, stability guarantees that the numerical method produces a bounded solution whenever the solution of the exact differential equation is bounded. Stability, in general, can be difficult to investigate, especially when the equation under consideration is nonlinear.

In certain cases, von Neumann stability is necessary and sufficient for stability in the sense of Lax–Richtmyer (as used in the Lax equivalence theorem): The PDE and the finite difference scheme models are linear; the PDE is constant-coefficient with periodic boundary conditions and have only two independent variables; and the scheme uses no more than two time levels.[4] Von Neumann stability is necessary in a much wider variety of cases. It is often used in place of a more detailed stability analysis to provide a good guess at the restrictions (if any) on the step sizes used in the scheme because of its relative simplicity.

Illustration of the method [edit]

The von Neumann method is based on the decomposition of the errors into Fourier series. To illustrate the procedure, consider the one-dimensional heat equation

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2}$$

defined on the spatial interval L, which can be discretized[5] as

$$(1) \qquad u_j^{n+1} = u_j^n + r\left(u_{j+1}^n - 2u_j^n + u_{j-1}^n\right)$$

where

$$r = \frac{\alpha\,\Delta t}{\Delta x^2}$$

and the solution u_j^n of the discrete equation approximates the analytical solution $u(x, t)$ of the PDE on the grid.

Define the round-off error ϵ_j^n as

$$\epsilon_j^n = N_j^n - u_j^n$$

where u_j^n is the solution of the discretized equation (1) that would be computed in the absence of round-off error, and N_j^n is the numerical solution obtained in finite precision arithmetic. Since the exact solution u_j^n must satisfy the discretized equation exactly, the error ϵ_j^n must also satisfy the discretized equation.[6] Thus

$$(2) \qquad \epsilon_j^{n+1} = \epsilon_j^n + r\left(\epsilon_{j+1}^n - 2\epsilon_j^n + \epsilon_{j-1}^n\right)$$

is a recurrence relation for the error. Equations (1) and (2) show that both the error and the numerical solution have the same growth or decay behavior with respect to time. For linear differential equations with periodic boundary condition, the spatial variation of error may be expanded in a finite Fourier series, in the interval L, as

$$(3) \qquad \epsilon(x) = \sum_{m=1}^{M} A_m e^{ik_m x}$$

where the wavenumber $k_m = \dfrac{\pi m}{L}$ with $m = 1, 2, \ldots, M$ and $M = L/\Delta x$. The time dependence of the error is included by assuming that the amplitude of error A_m is a function of time. Since the error tends to grow or decay exponentially with time, it is reasonable to assume that the amplitude varies exponentially with time; hence

$$(4) \qquad \epsilon(x,t) = \sum_{m=1}^{M} e^{at} e^{ik_m x}$$

where a is a constant.

Since the difference equation for error is linear (the behavior of each term of the series is the same as series itself), it is enough to consider the growth of error of a typical term:

$$(5) \qquad \epsilon_m(x,t) = e^{at} e^{ik_m x}$$

The stability characteristics can be studied using just this form for the error with no loss in generality. To find out how error varies in steps of time, substitute equation (5) into equation (2), after noting that

$$\epsilon_j^n = e^{at} e^{ik_m x}$$
$$\epsilon_j^{n+1} = e^{a(t+\Delta t)} e^{ik_m x}$$
$$\epsilon_{j+1}^n = e^{at} e^{ik_m (x+\Delta x)}$$
$$\epsilon_{j-1}^n = e^{at} e^{ik_m (x-\Delta x)},$$

to yield (after simplification)

$$(6) \qquad e^{a\Delta t} = 1 + \frac{a\Delta t}{\Delta x^2}\left(e^{ik_m \Delta x} + e^{-ik_m \Delta x} - 2\right).$$

Using the identities

$$\cos(k_m \Delta x) = \frac{e^{ik_m \Delta x} + e^{-ik_m \Delta x}}{2} \qquad \text{and} \qquad \sin^2 \frac{k_m \Delta x}{2} = \frac{1 - \cos(k_m \Delta x)}{2}$$

equation (6) may be written as

$$(7) \qquad e^{a\Delta t} = 1 - \frac{4a\Delta t}{\Delta x^2} \sin^2(k_m \Delta x/2)$$

Define the amplification factor

$$G \equiv \frac{\epsilon_j^{n+1}}{\epsilon_j^n} \qquad (8)$$

The necessary and sufficient condition for the error to remain bounded is that $|G| \leq 1$. However,

$$(8) \qquad G = \frac{e^{\alpha(t+\Delta t)} e^{ik_m x}}{e^{\alpha t} e^{ik_m x}} = e^{\alpha \Delta t}$$

Thus, from equations (7) and (8), the condition for stability is given by

$$(9) \qquad \left| 1 - \frac{4\alpha \Delta t}{\Delta x^2} \sin^2(k_m \Delta x/2) \right| \leq 1$$

Note that the term $\frac{4\alpha \Delta t}{\Delta x^2} \sin^2(k_m \Delta x/2)$ is always positive. Thus, to satisfy Equation (9):

$$(10) \qquad \frac{4\alpha \Delta t}{\Delta x^2} \sin^2(k_m \Delta x/2) \leq 2$$

For the above condition to hold at all $\sin^2(k_m \Delta x/2)$, we have

$$(11) \qquad \frac{\alpha \Delta t}{\Delta x^2} \leq \frac{1}{2}$$

Equation (11) gives the stability requirement for the FTCS scheme as applied to one-dimensional heat equation. It says that for a given Δx, the allowed value of Δt must be small enough to satisfy equation (10).

6 Well Posed PDE Problems

In the previous sections we saw some examples of partial differential equations. We now consider some important issues regarding the formulation and solvability of PDE problems. A solution to a PDE can be described as simply a function that reduces that PDE to an identity on some region of the independent variables. In general, a PDE alone, without any auxiliary boundary or initial conditions, will either have an infinity of solutions, or have no solution. Thus, in formulating a PDE problem there are three components: (i) the PDE; (ii) the region of space-time on which the PDE is required to be satisfied; (iii) the auxiliary boundary and initial conditions to be met.

For a PDE based mathematical model of a physical system to give useful results, it is generally necessary to formulate that model as what mathematicians call a well posed PDE problem. A PDE problem is said to be well posed if

1. a solution to the problem exists
2. the solution is unique, and
3. the solution depends continuously on the problem data.

(In a PDE problem the problem data consists of the coefficients in the PDE; the functions appearing in boundary and initial conditions; and the region on which the PDE is required to hold.)

التماسك :- consistancy

وهي من الخصائص الهامة لتقريبات الفروقة المنتهية (تقريبات محل المعادلات).

- ليس من الضروري أنه كلما تقاربت الفروقة المنتهية إلى المعادلات (معادلات التفاضلية) المقابلة حيث (لا يتقارب الحل للمعادلة التفاضلية) وذلك عندما تتعدى قياسات الفروقة في الشبكة إلى الصفر. وهنا تكون المجملة غير متماسكة.

نقول عن المجملة إنها متماسكة إذا صغنا تقريب المشتقات التفاضلية الجزئية في مجملة المعادلات التفاضلية الأصلية وذلك عندما $h, k \to 0$.

أي أنه هذا الاقتطاع يسعى إلى الصفر :-

$$\lim_{\substack{|h|\to 0 \\ |k|\to 0}} [f(u) - \tilde{f}(u)] \to 0 \iff \lim_{|h|,|k|\to 0} (T_\varepsilon) = 0$$

حيث: $f(u)$ المعادلة التفاضلية الجزئية.

$\tilde{f}(u)$ معادلة الفروقة. ، T_ε هذا اقتطاع المسألة.

عندئذ: هذا الاقتطاع $\tau_{ij}(u)$ في النقطة (ih, jk) يسعى للصفر بحيث يتحقق كالتالي :-

$$\tau_{ij}(u) = F_{ij}(u) - \tilde{F}_{ij}(u)$$

أي: $\tau_{ij} \to 0$ عندما $h \to 0$ ، $k \to 0$ نقول عن معادلة الفروقة إنها متماسكة.

مثال:

دراسة تماسك مجملة حسب شيم σ ، h ، k. (طريقة ضمنيا) :-

إن المعادلة : $\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0$ تتقرب بمعادلة الفروقة :-

$$\frac{w_{i,j+1} - w_{i,j}}{k} - \frac{\sigma w_{i+1,j+1} - 2(\sigma w_{i,j+1} + (1-\sigma) w_{i,j-1}) + w_{i-1,j}}{h}$$

وتملك هذا الاقتطاع :-

$$T_{ij} = \left(\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2}\right)_{ij} + \left[\frac{k^2}{6}\cdot\frac{\partial^3 u}{\partial t^3} - \frac{h^2}{12}\cdot\frac{\partial^4 u}{\partial x^4} + (2\sigma - 1)\frac{k}{h^2}\cdot\frac{\partial u}{\partial t} - \frac{k^2}{h^2}\frac{\partial^2 u}{\partial t^2}\right]_{ij}$$

$$+ O\left(\frac{k^3}{h^2}, h^4, k^4\right)$$

- في حالة: $k = \lambda . h$ عندما $h \to 0$ فناقش حالات σ كالتالي :-

Fig

- $\sigma \neq \frac{1}{2}$ عندئذ :-

$$\lim_{h\to 0} (2\sigma - 1)\frac{\lambda h}{h^2}\cdot\frac{\partial u}{\partial t} - \frac{\lambda^2 h^2}{h^2}\cdot\frac{\partial^2 u}{\partial t^2} \to \infty$$

- $\sigma = \frac{1}{2}$ عندئذ :-

$$\lim_{h\to 0} \frac{\lambda^2 h^2}{h^2}\cdot\frac{\partial^2 u}{\partial t^2} \neq 0$$

وبالتالي المجملة في هذه الحالة غير متماسكة.

- في حالة $k = \lambda . h^2$ عندما $h \to 0$ أيضا نناقش حالات σ كما يلي :-

- عندئذٍ: $\sigma = \frac{1}{2}$

$$\lim_{h \to 0} \frac{(\lambda h^2)^2}{h^2} \cdot \frac{\partial^2 u}{\partial t^2} = 0 \qquad \left(\frac{\lambda^2 h^4}{h^2} \cdot \lambda^2 h^2 \xrightarrow[h \to 0]{} 0 \right)$$

- عندئذٍ: $\sigma \neq \frac{1}{2}$

$$\lim_{h \to 0} \left[(2\sigma - 1) \cdot \frac{\lambda h^2}{h^2} \cdot \frac{\partial u}{\partial t} - \frac{\lambda^2 h^4}{h^2} \cdot \frac{\partial^2 u}{\partial t^2} \right] \neq 0$$

وبالتالي المعادلة غير متماسكة.

وهو المطلوب

مثال: لتكن $f(u) = u_x + v \cdot u_y$ حيث v ثابت.

ومعادلة الفروق المنتهية لها هي من الشكل:

$$\tilde{f}(u) = \frac{u_{i+1,j} - u_{i,j}}{h} + v \cdot \frac{u_{i,j+1} - u_{i,j}}{h}$$

حيث: $h = h_1 = h_2$ (لأجل التغير في الشبكة).

إنّ منشور تايلور حول نقاط الشبكة $(x = ih, y = ih)$ توضح تماسك المعادلة كالتالي:

$$u_{i+1,j} = u_{i,j} + h \cdot u_x + \frac{h^2}{2} \cdot u_{xx} + O(h^3)$$

$$u_{i,j+1} = u_{i,j} + h \cdot u_y + \frac{h^2}{2} \cdot u_{yy} + O(h^3)$$

$$f(u) - \tilde{f}(u) = -\frac{h}{2}(u_{xx} + v \cdot u_{yy}) + O(h^2) \xrightarrow[h \to 0]{} 0$$

وبالتالي المعادلة متماسكة.

وهو المطلوب

التقارب والاستقرار :- convergence and stability

- نقصد بالتقارب أنه نتائج الطريقة تتقارب نحو الحل التحليلي عندما h , k تؤول نحو الصفر.
- والاستقرار هو أنه الخطأ في إحدى مراحل الحسابات لا يؤدي إلى نمو الخطأ بالتقريب المتكرر.

وذلك باعتبار الخطأ هو الفرق بين الحل العددي (التقريبي) والحل الفعلي للمعادلة.

لتكن المعادلة :-

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \quad ; \quad 0 < x < 1 \quad ; \quad t > 0$$

ليكن u_{ij} هو الحل الفعلي , w_{ij} هو الحل التقريبي للمعادلة (معادلة الفروقة) وباعتبار الخطأ هو الفرق بين الحلين e_{ij} .

مثال:

الفروقة المستهية القريبة الظاهرية :-

$$\frac{w_{i,j+1} - w_{ij}}{k} = \frac{w_{i-1,j} - 2w_{ij} + w_{i+1,j}}{h^2}$$

في نقاط الشبكة :-

$$w_{ij} = u_{ij} - e_{ij} \quad , \quad w_{ij+1} = u_{ij+1} - e_{ij+1}$$

نعوضه في معادلة الفروقة :-

$$e_{ij+1} = \lambda e_{i-1j} + (1-2\lambda) e_{ij} + \lambda e_{i+1j} + u_{ij+1} - u_{ij} + \lambda(2u_{ij} - u_{i-1j} - u_{i+1j})$$

باستخدام تايلور نحصل على :-

$$u_{i+1j} = u(x_i + h, t_j) = u_{ij} + h^2\left(\frac{\partial u}{\partial x}\right)_{ij} + \frac{h^2}{2} \cdot \frac{\partial^2 u}{\partial x^2}(x_i + \theta_1 h, t_j)$$

$$u_{i-1j} = u(x_i - h, t_j) = u_{ij} - h^2\left(\frac{\partial u}{\partial x}\right)_{ij} + \frac{h^2}{2} \cdot \frac{\partial^2 u}{\partial x^2}(x_i - \theta_2 h, t_j)$$

$$u_{ij+1} = u(x_i, t_j + k) = u_{ij} + k \cdot \frac{\partial u}{\partial x}(x_i, t_j + \theta_3 k)$$

حيث :-

$$0 < \theta_1, \theta_2, \theta_3 < 1$$

بالتعويض في المعادلة الأصلية نحصل على :-

$$e_{ij+1} = \lambda e_{i-1j} + (1-2\lambda) e_{ij} + \lambda e_{i+1j} + \underbrace{\left[\frac{\partial u}{\partial x}(x_i, t_j + \theta_3 k) - \frac{\partial^2 u}{\partial x^2}(x_i - \theta_4 h, t_j)\right]}_{[*]}$$

حيث :-

$$-1 < \theta_3, \theta_4 < 1$$

وهذه هي معادلة الفروقة للخطأ e_{ij} والتي لا نحتاج لحلها (لدراسة الاستقرار).

لتكن E_j القيمة المطلقة للخطأ على طول السطر j , M هي max لعبارة كبيرة لكمية $[*]$ عندما $r \leq \frac{1}{2}$ كل معاملات e في المعادلة إما موجبة أو صفر وبالتالي :-

$$|e_{ij+1}| \leq \lambda |e_{i-1j}| + (1-2\lambda)|e_{ij}| + \lambda |e_{i+1j}| + kM$$

$$\leq \lambda E_j - (1-2\lambda) E_j + \lambda E_j + kM$$

$$= E_j + kM$$

وأيضاً :-

(شرط الاستقرار)

$$E_{j+1} \leq E_j + kM \leq (E_{j-1} + kM) + kM \leq \dots \leq E_0 + (j+1)kM = t_{j+1} M$$

وبما أننا نتعامل مع الشروط الحدية غير الغالبة للاستقامة $E_0 = 0$.

عندما $h \to 0$ فإنه $k = r h^2 \to 0$ و M ستعتمد :- $\left(\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2}\right)_{ij}$ (الوهو شرط الذي نطبقه في الحل الفعلي والتقريبي) .

ولطالما أنه الطريقة الحدية متماسكة عند قيمة M لذلك $E_{j+1} \to 0$ كما أنه :

$$|u_{ij} - w_{ij}| \leq E_i$$

وبالتالي سنتوقع أنه نقول أنه $w \to u$ عندما $h \to 0$ حيث $r \leq \frac{1}{2}$

أي الحلول متقاربة (مستقرة ومتماسكة).

عندما $r > \frac{1}{2}$ يمكننا أن نظهر أنه الحلول تتعدى إلى ∞ عندما $h \to 0$ (متباعدة).

خلاصة :

توضح لنا من خلال دراسة التماسك والاستقرار النقاط البسيطة التالي :-

عند دراسة التماسك يكون الانتقال بالتالي من

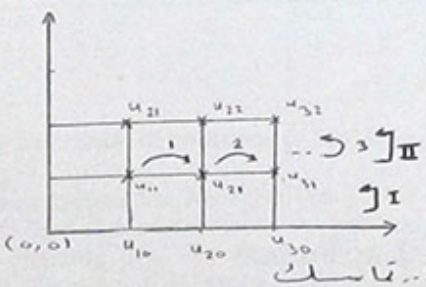
النقطة إلى التي تليها في لحظة معينة بعد المرور بالنقط

جميعها تنتقل إلى اللحظة التي تليها وهكذا .

وعند دراسة الاستقرار سنأخذ في لحظة معينة

نأخذ الخطأ الأعظم في تلك اللحظة وننتقل إلى

اللحظة التي تليها .



1-2-3... تماسك

I - II ... استقرار .

التماسك + الاستقرار $\Leftrightarrow$ التقارب .

convergence $\Leftrightarrow$ stability + consistancy .

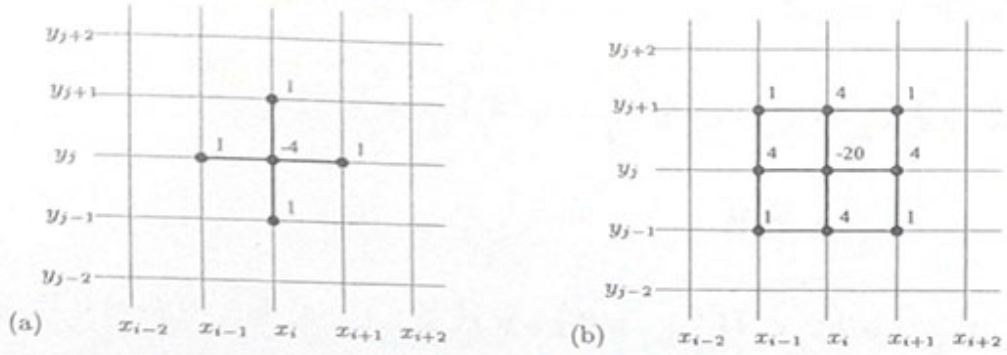


Figure 3.1: Portion of the computational grid for a two-dimensional elliptic equation. (a) The 5-point stencil for the Laplacian about the point (i, j) is also indicated. (b) The 9-point stencil is indicated, which is discussed in Section 3.4.

We also need to specify boundary conditions all around the boundary of the region Ω. These could be Dirichlet conditions, where the temperature $u(x, y)$ is specified at each point on the boundary, or Neumann conditions, where the normal derivative (the heat flux) is specified. We may have Dirichlet conditions specified at some points on the boundary and Neumann conditions at other points.

In one space dimension the corresponding Laplace's equation $u''(x) = 0$ is trivial: the solution is a linear function connecting the two boundary values. In two dimensions even this simple equation in nontrivial to solve, since boundary values can now be specified at every point along the curve defining the boundary. Solutions to Laplace's equation are called *harmonic functions*. You may recall from complex analysis that if $g(z)$ is any complex analytic function of $z = x + iy$, then the real and imaginary parts of this function are harmonic. For example, $g(z) = z^2 = (x^2 - y^2) + 2ixy$ is analytic and the functions $x^2 - y^2$ and $2xy$ are both harmonic.

The operator ∇^2 defined by

$$\nabla^2 u = u_{xx} + u_{yy}$$

is called the *Laplacian*. The notation ∇^2 comes from the fact that, more generally,

$$(\kappa u_x)_x + (\kappa u_y)_y = \nabla \cdot (\kappa \nabla u)$$

where ∇u is the gradient of u,

$$\nabla u = \begin{bmatrix} u_x \\ u_y \end{bmatrix} \tag{3.7}$$

and $\nabla \cdot$ is the divergence operator,

$$\nabla \cdot \begin{bmatrix} u \\ v \end{bmatrix} = u_x + v_y. \tag{3.8}$$

The symbol Δ is also often used for the Laplacian, but would lead to confusion in numerical work where Δx and Δy will be used for grid spacing.

3.2 The five-point stencil for the Laplacian

To discuss discretizations, first consider the Poisson problem (3.5) on the unit square $0 \le x \le 1$, $0 \le y \le 1$ and suppose we have Dirichlet boundary conditions. We will use a uniform Cartesian grid consisting of grid points (x_i, y_j) where $x_i = i\Delta x$ and $y_j = j\Delta y$. A section of such a grid is shown in Figure 3.1.

Let u_{ij} represent an approximation to $u(x_i, y_j)$. In order to discretize (3.5) we replace both the x- and y-derivatives with centered finite differences, which gives

$$\frac{1}{(\Delta x)^2}(u_{i-1,j} - 2u_{ij} + u_{i+1,j}) + \frac{1}{(\Delta y)^2}(u_{i,j-1} - 2u_{ij} + u_{i,j+1}) = f_{ij}. \tag{3.9}$$

For simplicity of notation we will consider the special case where $\Delta x = \Delta y \equiv h$, though it is easy to handle the general case. We can then rewrite (3.9) as

$$\frac{1}{h^2}(u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1} - 4u_{ij}) = f_{ij}. \tag{3.10}$$

This finite difference scheme can be represented by the *5-point stencil* shown in Figure 3.1. We have both an unknown u_{ij} and an equation of the form (3.10) at each of m^2 grid points for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, m$, where $h = 1/(m+1)$ as in one dimension. We thus have a linear system of m^2 unknowns. The difference equations at points near the boundary will of course involve the known boundary values, just as in the one-dimensional case, which can be moved to the right-hand side.

If we collect all of these equations together into a matrix equation, we will have an $m^2 \times m^2$ matrix which is very *sparse*, *i.e.*, most of the elements are zero. Since each equation involves at most 5 unknowns (less near the boundary), each row the matrix has at most 5 nonzeros and at least $m^2 - 5$ elements that are zero. This is analogous to the tridiagonal matrix (2.9) seen in the one-dimensional case, in which each row has at most 3 nonzeros.

Unfortunately, in two space dimensions the structure of the matrix is not as simple as in one dimension, and in particular the nonzeros will not all be as nicely clustered near the main diagonal. The exact structure of the matrix depends on the order in which we order the unknowns and write down the equations, as we will see below, but no ordering is ideal.

Note that in general we are always free to change the order of the equations in a linear system without changing the solution. Modifying the order corresponds to permuting the rows of the matrix and right-hand side. We are also free to change the ordering of the unknowns in the vector of unknowns, which corresponds to permuting the columns of the matrix. As an example, consider the one-dimensional difference equations given by (2.9). Suppose we reordered the unknowns by listing first the unknowns at odd numbered grid points and then the unknowns at even numbered grid points, so that $\tilde{U} = [U_1, U_3, U_5, \ldots, U_2, U_4, \ldots]^T$. If we also reorder the equations in the same way, *i.e.*, we write down first the difference equation centered at U_1, then at U_3, U_5, etc., then we would obtain the following system:

$$\frac{1}{h^2}\left[\begin{array}{ccccccc|ccccccc}
-2 & & & & & & & 1 & & & & & & \\
 & -2 & & & & & & 1 & 1 & & & & & \\
 & & -2 & & & & & & 1 & 1 & & & & \\
 & & & \ddots & & & & & & \ddots & \ddots & & & \\
 & & & & & & -2 & & & & & & 1 & 1 \\
\hline
1 & 1 & & & & & & -2 & & & & & & \\
 & 1 & 1 & & & & & & -2 & & & & & \\
 & & 1 & 1 & & & & & & -2 & & & & \\
 & & & \ddots & \ddots & & & & & & \ddots & & & \\
 & & & & & & 1 & & & & & & & -2
\end{array}\right]
\left[\begin{array}{c} U_1 \\ U_3 \\ U_5 \\ \vdots \\ U_{m-1} \\ \hline U_2 \\ U_4 \\ U_6 \\ \vdots \\ U_m \end{array}\right]
=
\left[\begin{array}{c} f(x_1) - \alpha/h^2 \\ f(x_3) \\ f(x_5) \\ \vdots \\ f(x_{m-1}) \\ \hline f(x_2) \\ f(x_4) \\ f(x_6) \\ \vdots \\ f(x_m) - \beta/h^2 \end{array}\right]. \tag{3.11}$$

This linear system has the same solution as (2.9) modulo the reordering of unknowns, but looks very different. For this one-dimensional problem there is no point in reordering things this way, and the natural ordering $[U_1, U_2, U_3, \ldots]^T$ clearly gives the optimal matrix structure for the purpose of applying Gaussian elimination. By ordering the unknowns so that those which occur in the same equation are close to one another in the vector, we keep the nonzeros in the matrix clustered near the diagonal.

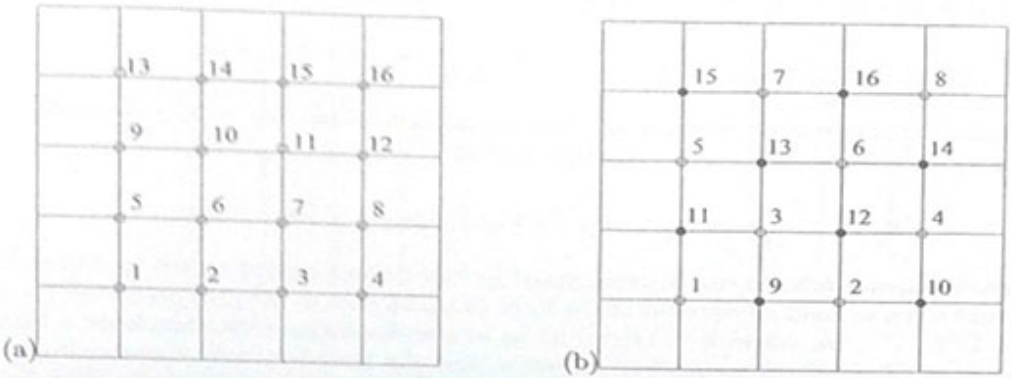


Figure 3.2: (a) The natural rowwise order of unknowns and equations on a 4×4 grid. (b) The red-black ordering.

Returning to the two-dimensional problem, it should be clear that there is no way to order the unknowns so that all nonzeros are clustered adjacent to the diagonal. About the best we can do is to use the *natural rowwise ordering*, where we take the unknowns along the bottom row, $u_{11}, u_{21}, u_{31}, \ldots, u_{m1}$, followed by the unknowns in the second row, $u_{12}, u_{22}, \ldots, u_{m2}$, and so on, as illustrated in Figure 3.2(a). This gives a matrix equation where A has the form

$$A = \frac{1}{h^2}\begin{bmatrix} T & I & & & \\ I & T & I & & \\ & I & T & I & \\ & & \ddots & \ddots & \ddots \\ & & & I & T \end{bmatrix} \tag{3.12}$$

which is an $m \times m$ *block tridiagonal matrix* in which each block T or I is itself an $m \times m$ matrix,

$$T = \begin{bmatrix} -4 & 1 & & & \\ 1 & -4 & 1 & & \\ & 1 & -4 & 1 & \\ & & \ddots & \ddots & \ddots \\ & & & 1 & -4 \end{bmatrix}$$

and I is the $m \times m$ identity matrix. While this has a nice structure, the 1 values in the I matrices are separated from the diagonal by $m-1$ zeros, since these coefficients correspond to grid points lying above or below the central point in the stencil and hence are in the next or previous row of unknowns.

Another possibility, which has some advantages in the context of certain iterative methods, is to use the *red-black ordering* (or checkerboard ordering) shown in Figure 3.2. This is the two-dimensional analog of the odd-even ordering that leads to the matrix (3.11) in one dimension. This ordering is significant because all 4 neighbors of a red grid point are black points, and vice versa, and leads to a matrix equation with the structure

$$\begin{bmatrix} D & H \\ H^T & D \end{bmatrix} \begin{bmatrix} u_{\text{red}} \\ u_{\text{black}} \end{bmatrix} = \cdots \tag{3.13}$$

where $D = -\frac{4}{h^2} I$ is a diagonal matrix of dimension $m^2/2$ and H is determined in Exercise 3.1.

3.3 Accuracy and stability

The discretization of the two-dimensional Poisson problem can be analyzed using exactly the same approach as we used for the one-dimensional boundary value problem. The local truncation error τ_{ij} at the (i, j) grid point is defined in the obvious way,

$$\tau_{ij} = \frac{1}{h^2}(u(x_{i-1}, y_j) + u(x_{i+1}, y_j) + u(x_i, y_{j-1}) + u(x_i, y_{j+1}) - 4u(x_i, y_j)) - f(x_i, y_j),$$

and by splitting this into the second order difference in the x- and y-directions it is clear from previous results that

$$\tau_{ij} = \frac{1}{12}h^2(u_{xxxx} + u_{yyyy}) + O(h^4).$$

For this linear system of equations the global error $E_{ij} = u_{ij} - u(x_i, y_j)$ then solves the linear system

$$A^h E^h = -\tau^h$$

just as in one dimension, where A^h is now the discretization matrix with mesh spacing h, e.g., the matrix (3.12) if the rowwise ordering is used. The method will be globally second order accurate in some norm provided that it is stable, *i.e.*, that $\|(A^h)^{-1}\|$ is uniformly bounded as $h \to 0$.

In the 2-norm this is again easy to check, for this simple problem, since we can explicitly compute the spectral radius of the matrix, as we did in one dimension in Section 2.10. The eigenvalues and eigenvectors of A can now be indexed by 2 parameters p and k corresponding to wavenumbers in the x and y directions, for $p, k = 1, 2, \ldots, m$. The (p, k) eigenvector $u^{p,k}$ has the m^2 elements

$$u_{ij}^{p,k} = \sin(p\pi ih)\sin(k\pi jh). \tag{3.14}$$

The corresponding eigenvalue is

$$\lambda_{p,k} = \frac{2}{h^2}\left((\cos(p\pi h) - 1) + (\cos(k\pi h) - 1)\right). \tag{3.15}$$

The eigenvalues are strictly negative (A is negative definite) and the one closest to the origin is

$$\lambda_{1,1} = -2\pi^2 + O(h^2).$$

The spectral radius of $(A^h)^{-1}$, which is also the 2-norm, is thus

$$\rho((A^h)^{-1}) = 1/\lambda_{1,1} \approx -1/2\pi^2$$

Hence the method is stable in the 2-norm.

While we are at it, let's also compute the condition number of the matrix A^h, since it turns out that this is a critical quantity in determining how rapidly certain iterative methods converge. Recall that the 2-norm condition number is defined by

$$\kappa_2(A) = \|A\|_2\|A^{-1}\|_2.$$

We've just seen that $\|(A^h)^{-1}\|_2 \approx -1/2\pi^2$ for small h, and the norm of A is given by its spectral radius. The largest eigenvalue of A (in magnitude) is

$$\lambda_{m,m} \approx -\frac{8}{h^2}$$

and so

$$\kappa_2(A) \approx \frac{2}{\pi^2 h^2} = O(1/h^2) \quad \text{as } h \to 0. \tag{3.16}$$

The fact that the matrix becomes more ill-conditioned as we refine the grid is responsible for the slow-down of iterative methods, as discussed in Chapter 5.

(16)