1. You use three different iterative methods to find the minimum of the function

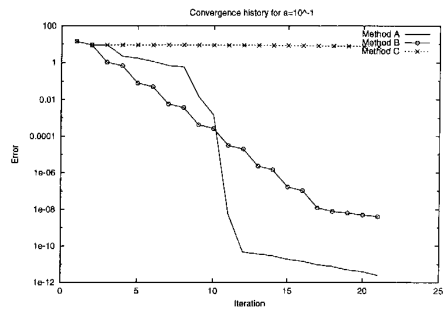
$$f(x, y) = -\frac{1}{1 + x^2 + a y^2},$$

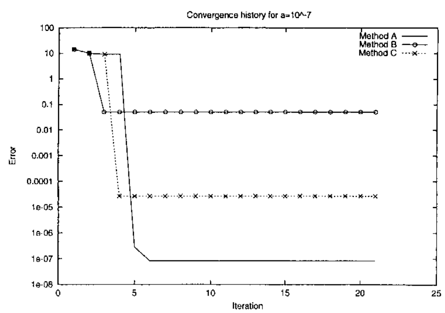
B → (a) The gradient method;

C → (b) The Fletcher–Reeves conjugate gradient method without restart;

A → (c) The Fletcher–Reeves conjugate gradient method with restart every second iteration.

The following two graphs show the decrease of the error with the number of iterations for two different values of a.





Match the method used to the graphs shown (the labeling is the same in both plots!), and explain your choice. (10)

Key features:

- The gradient method converges slowly but reliably for a problem with moderate condition number. (case $a = 10^{-1}$)
- The gradient method converges poorly (large residual error) when the condition number is large. (case $a = 10^{-7}$)
- CG with restart does best overall, although initial convergence is not as fast as for the gradient method.
- Note that the graphs of CG with/without restart are identical for two iterations
- CG without restart can converge very poorly (case $a = 1$), as it always keeps past information, and thus never "forgets" an initially "wrong" orthonormality relation.

Remark: When using CG with the Pollack-Ribière formula, CG without restart does much better and matches, or sometimes slightly exceeds, the convergence rate of CG with restart.

2. Show that the stochastic differential equation

$$dX = \tfrac{1}{3} X^{1/3}\,dt + X^{2/3}\,dW, \qquad X(0) = X_0,$$

is solved by

$$X(t) = \left(X_0^{1/3} + \tfrac{1}{3} W(t)\right)^3 .$$

(10)

Use the Itô-formula

$$d f(t,W) = \frac{\partial f}{\partial t}\,dt + \frac{\partial f}{\partial W}\,dW + \frac{1}{2}\frac{\partial^2 f}{\partial W^2}\,dt$$

$$\Rightarrow \quad dX = 3\left(X_0^{1/3} + \tfrac{1}{3}W(t)\right)^2 \cdot \tfrac{1}{3}\,dW + \tfrac{1}{2}\cdot 2\underbrace{\left(X_0^{1/3} + \tfrac{1}{3}W(t)\right)}_{= X^{1/3}} \cdot \tfrac{1}{3}\,dt$$

$$= \tfrac{1}{3} X^{\frac{1}{3}}\,dt + X^{\frac{2}{3}}\,dW$$

3. Apply the Euler–Maruyama method

$$X_{j+1} = X_j + f(X_j)\,\Delta t + g(X_j)\,\Delta W_j$$

to the stochastic differential equation

$$dX = \mu X\,dW .$$

Show that

$$\mathbb{E}[|X_{j+1}|^2] = (1 + \Delta t\,|\mu|^2)\,\mathbb{E}[|X_j|^2] . \qquad (*)$$

(10)

Here $X_{j+1} = X_j + \mu X_j \Delta W_j = X_j (1 + \mu \Delta W_j)$

$$\Rightarrow \quad |X_{j+1}|^2 = |X_j|^2 \underbrace{|1 + \mu \Delta W_j|^2}_{= 1 + 2\,\mathrm{Re}\,\mu\,\Delta W_j + |\mu|^2 \Delta W_j^2}$$

$$\Rightarrow \quad \mathbb{E}[|X_{j+1}|^2] = \mathbb{E}[|X_j|^2] + 2\,\mathrm{Re}\,\mu\,\mathbb{E}[|X_j|^2 \Delta W_j] + |\mu|^2\,\mathbb{E}[|X_j|^2 \Delta W_j^2]$$

Since X_j and ΔW_j are statistically independent (X_j depends only on the past, and ΔW_j is the new increment),

$$\mathbb{E}[|X_j|^2 \Delta W_j] = \mathbb{E}[|X_j|^2]\,\underbrace{\mathbb{E}[\Delta W_j]}_{=0} = 0,$$

$$\mathbb{E}[|X_j|^2 \Delta W_j^2] = \mathbb{E}[|X_j|^2]\,\mathbb{E}[\Delta W_j^2] = \Delta t\,\mathbb{E}[|X_j|^2]$$

The result (*) directly follows.

4. On the interval $[0, 2]$, consider the boundary value problem

$$-y''(x) = f(x)\,, \quad y'(0) = y'(2) = 0\,. \qquad (*)$$

(a) Away from the boundary, the solution is approximated by

$$-\frac{y_{j-1} - 2y_j + y_{j+1}}{h^2} = f_j\,. \qquad (**)$$

Show that the local truncation error of this method is of order 2.

(b) We approximate the boundary conditions by

$$\frac{y_1 - y_0}{h} = 0\,,$$

with a corresponding expression for the second boundary condition. Show that this approximation is accurate only to order 1.

(c) Suggest an improvement that ensures the method is of order 2 up to the boundary.

(d) Write out the resulting system of linear equations of your method, or of the method given in (a) and (b), with only three nodes $x_0 = 0$, $x_1 = 1$, and $x_2 = 2$. Is the resulting matrix invertible?

(e) Do you expect the solution to the original problem to be unique? (Think of what happens to constants...)

(f) **Extra credit:** Can you think of a reasonable condition that would make the solution unique? Note that you must not add a third boundary condition, as it would generically overdetermine the system.

(5+5+5+5+5+*10*)

(a)
$$y(x-h) = y(x) - y'(x)h + \tfrac{1}{2}y''(x)h^2 - \tfrac{1}{3!}y'''(x)h^3 + O(h^4)$$
$$-2y(x) = -2y(x)$$
$$+)\quad y(x+h) = y(x) + y'(x)h + \tfrac{1}{2}y''(x)h^2 + \tfrac{1}{3!}y'''(x)h^3 + O(h^4)$$

$$y(x-h) - 2y(x) + y(x+h) = y''(x)h^2 + O(h^4)$$

Plugging in $y(x-h)$ for y_{j-1}, etc., why y solves $(*)$, gives

$$T_j = -y''(x_j) + O(h^2) - f(x_j) = O(h^2)\,.$$

(b) $$y(0+h) = y(0) + y'(0)h + \tfrac{1}{2}y''(0)h^2 + O(h^3)$$

$\Rightarrow$ At the boundary, the local truncation error is

$$T^{\text{bound}} = \frac{y(0+h) - y(0)}{h} = \underbrace{y'(0)}_{=0} + \tfrac{1}{2}y''(0)h + O(h^2)$$
$$= O(h)$$

(c) Solution 1:

For $j = 1, \dots, n-1$, use formula $(**)$.

For $j = 0$, write

$$\frac{y_1 - y_0}{h} + \frac{1}{2}f_0 h = 0\,,$$

killing the $O(h)$ term of T^{bound}, with a similar expression at $j = n$.

Solution 2:

At $j = 0$, use a centered difference, which can easily shown to be second order,

$$\frac{y_1 - y_{-1}}{2h} = 0\,. \qquad (***)$$

we eliminate the fictitious grid point y_{-1} by using (**) with $j=0$:

$$-\frac{y_{-1} - 2y_0 + y_1}{h^2} = f_0$$

$$\Rightarrow \quad y_{-1} = -h^2 f_0 + 2y_0 - y_1$$

Plugging back into (***), we obtain

$$0 = \frac{y_1 - \left(-h^2 f_0 + 2y_0 - y_1\right)}{2h} = \frac{y_1 - y_0}{h} + \frac{1}{2} f_0 h$$

I.e., we find the same expression as for solution 1 above.

(d)
$$\frac{y_1 - y_0}{h} = -\frac{1}{2} f_0 h$$

$$\frac{y_0 - 2y_1 + y_2}{h^2} = -f_1$$

$$\frac{y_2 - y_1}{h} = \frac{1}{2} f_2 h$$

Here, $h=1$, so that we obtain the matrix equation

$$\begin{pmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} f_0 \\ -f_1 \\ -\frac{1}{2} f_2 \end{pmatrix}$$

The matrix on the left is clearly singular, so the system is only solvable if the right side is compatible; if a solution exists, it is not unique.

(e) Replacing $y \to y + c$ for an ~~arbt~~ constant c does not change the equation, nor does it change the boundary conditions. Thus, we expect (at least) a one-parameter family of solutions (if they exist).

(f) It can be shown that the solution is indeed unique up to addition of constant:

Assume that $y(x)$ and $z(x)$ both solve (*). Set

$$w = y - z$$

and compute

$$\int_0^2 (w')^2 \, dx \overset{\text{I.b.p.}}{=} \underbrace{w(x)\, w'(x) \Big|_0^2}_{\substack{=0 \text{ because } w \\ \text{must also satisfy the} \\ \text{boundary condition}}} - \int_0^2 \underbrace{w''}_{\substack{= y'' - z'' \\ = -f - (-f) \\ = 0}} w \, dx$$

$$= 0.$$

But $\int_0^2 (w')^2 \, dx = 0$ only if $w' = 0$,

i.e. if y and z differ by a constant.

Thus, the solution can be made unique by any additional condition which fixes the constant.
It is common to prescribe the mean, e.g.

$$\int_0^2 y(x)\, dx \equiv 0 .$$

Remark: Since the solution to the "unconstrained" problem is not unique, it is — as in the discrete case — solvable only if f is "compatible", namely if

$$\int_0^2 f(x)\, dx = -\int_0^2 y''(x)\, dx \overset{\text{F.T.C.}}{=} -\left(y'(2) - y'(0)\right) = 0 .$$

5. Recall that the Householder reflector about the hyperplane normal to v is the matrix

$$\mathrm{H} = \mathrm{I} - 2\frac{vv^T}{v^Tv}.$$

(a) What are the eigenvalues of H?
Hint: What do you get by applying H to v, or to a vector orthogonal to v?

(b) Find v and α such that

$$\mathrm{H}\underbrace{\begin{pmatrix}0\\4\\3\\0\end{pmatrix}}_{=:x} = \underbrace{\begin{pmatrix}\alpha\\0\\0\\0\end{pmatrix}}_{=:y}.$$

(10+10)

(a) $$Hv = v - 2\frac{v\,v^Tv}{v^Tv} = -v$$

$\Rightarrow$ -1 is an eigenvalue with eigenvector v.

If w is orthogonal to v,

$$Hw = w - 2\frac{v\,v^Tw}{v^Tv} = w$$

$\Rightarrow$ 1 is an eigenvector whose $(n-1)$ dimensional eigenspace is the hyperplane normal to v.

(b) We are looking for a vector x s.t.

$$Hx = y$$

$$\Rightarrow \quad x - 2\frac{v\,v^Tx}{v^Tv} = y$$

$$\Rightarrow \quad x - y = 2v \frac{v^T x}{v^T v}$$

We see immediately that v must be a scalar multiple of $x-y$, and since the expression is invariant under scaling of v, we may choose $v = x-y$ without loss of generality.

$$\Rightarrow \quad v = 2v \frac{v^T x}{v^T v}$$

$$\Rightarrow \quad 0 = v\left(2\frac{v^T x}{v^T v} - 1\right) = \frac{v}{v^T v}\left(2v^T x - v^T v\right)$$

We thus demand that

$$2v^T x = v^T v$$

$$\Rightarrow 2(x-y)^T x = (x-y)^T (x-y)$$

$$\Rightarrow 2x^T x - 2y^T x = x^T x - 2x^T y + y^T y$$

$$\Rightarrow x^T x = y^T y = \alpha^2$$

<u>Note:</u> It is not required to derive this expression on the exam. You could have remembered the expression and checked by direct computation.

Here: $\quad x^T x = 3^2 + 4^2 = 25 = \alpha^2$

we choose $\alpha = 5$, so that

$$v = x - y = \begin{pmatrix} 0 \\ 4 \\ 3 \\ 0 \end{pmatrix} - \begin{pmatrix} 5 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -5 \\ 4 \\ 3 \\ 0 \end{pmatrix}$$

Direct check (not required):

$$v^T v = 3^2 + 4^2 + 5^2 = 50$$

$$v v^T = \begin{pmatrix} 25 & -20 & -15 & 0 \\ -20 & 16 & 12 & 0 \\ -15 & 12 & 9 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\Rightarrow H = I - \frac{1}{25} v v^T$$

$$Hx = \begin{pmatrix} 0 \\ 4 \\ 3 \\ 0 \end{pmatrix} - \frac{1}{25} \underbrace{\begin{pmatrix} 25 & -20 & -15 & 0 \\ -20 & 16 & 12 & 0 \\ -15 & 12 & 9 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 4 \\ 3 \\ 0 \end{pmatrix}}_{= \begin{pmatrix} -125 \\ 100 \\ 75 \\ 0 \end{pmatrix}} = \begin{pmatrix} 5 \\ 0 \\ 0 \\ 0 \end{pmatrix} \checkmark$$

6. You minimize the function

$$f(x, y) = x^2 - y^2$$

where x and y are constrained to the unit circle, i.e.

$$h(x, y) = x^2 + y^2 - 1 = 0 .$$

(a) State the exact solution to this problem.
(No computation required, the problem is simple enough to spot the answer.)

(b) Solve the problem using the quadratic penalty method, i.e. minimize

$$p_\alpha(x, y) = f(x, y) + \alpha h^2(x, y) .$$

Compute the minimizer (x^*_α, y^*_α) of the penalized problem explicitly.

(c) Show that $h(x^*_\alpha, y^*_\alpha) \to 0$ as $\alpha \to \infty$.

(10+10+5)

(a) Clearly f decreases fastest along the y-axis, so the minimum is at its intersection with the unit circle,

$$(0, \pm 1)$$

(b) $p_\alpha = x^2 - y^2 + \alpha^2 h^2$

$$\Rightarrow \nabla p_\alpha = \begin{pmatrix} 2x + 2\alpha h \frac{\partial h}{\partial x} \\ -2y + 2\alpha h \frac{\partial h}{\partial y} \end{pmatrix} = 2 \begin{pmatrix} x + \alpha h \, 2x \\ -y + \alpha h \, 2y \end{pmatrix} \equiv 0$$

$\Rightarrow \quad x(1 + 2\alpha h) = 0$

and $y(-1 + 2\alpha h) = 0$

We have to distinguish four cases

i) $x = y = 0$: for α large a maximum $\Rightarrow$ discarded

(ii) $1 + 2\alpha h = 0$ and $-1 + 2\alpha h = 0$: inconsistent $\Rightarrow$ discarded

(iii) $y = 0$ and $1 + 2\alpha h = 0$: Clearly larger value for p_α than with the fourth alternative $\Rightarrow$ discarded

(iv) $x = 0$ and $-1 + 2\alpha h = 0$ good!

$\Rightarrow \quad x^*_\alpha = 0$ and $1 = 2\alpha h(x^*_\alpha, y^*_\alpha)$

$$\Rightarrow 1 = 2\alpha \left((y^*_\alpha)^2 - 1 \right)$$

$$\Rightarrow y^*_\alpha = \pm \sqrt{1 - \frac{1}{2\alpha}}$$

(c) Clearly, $y^*_\alpha \to \pm 1$ as $\alpha \to \infty$

Since $h(x,y)$ is continuous, $h(x^*_\alpha, y^*_\alpha) \to 0^2 + 1^2 - 1 = 0$ as $\alpha \to \infty$