Solutions

1 a) Denote $r = Ax - b$; $x = \arg\min\|r\|_2 = \arg\min r^T r = \arg\min(x^T A^T Ax - 2(A^T b)^T x + b^T b)$

Differentiating with respect to x and setting the result to zero, we get $2A^T Ax - 2A^T b = 0$, or

$$A^T Ax = A^T b \qquad (2)$$

A is of full column rank => $A^T A$ is nonsingular => (2) has a unique solution.

$A^T A$ is symmetric and positive definite => we can compute its Cholesky factorization

$$A^T A = LL^T,$$

where L - lower triangular. This allows us to reduce (2) to solving the triangular systems

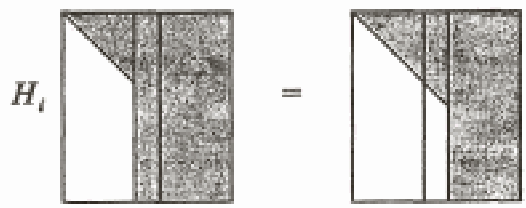
$$Ly = A^T b, \qquad L^T x = y.$$

b)First we want to decompose A into the product $Q\begin{bmatrix} R \\ 0 \end{bmatrix}$, $Q : m \times m$ orthogonal, $R : n \times n$ upper-triangular; since A is of full column rank, R is nonsingular.

To reduce A to the upper-triangular form, we successively apply Householder transformations

$H_n \ldots H_1 A = \begin{bmatrix} R \\ 0 \end{bmatrix}$, where $H_i = I - 2\dfrac{v_i v_i^T}{v_i^T v_i}$ => orthogonal and symmetric.

H_i reflects a vector against the hyperplane $v_i^{\perp}$. We choose v_i so that H_i zeros out the subdiagonal part of the ith column $\tilde{a}_i$ of the current state of A - $(H_{i-1} \ldots H_1 A)$



If $\tilde{a}_i' = \{\tilde{a}_i$ with the upper i-1 entries set to 0$\}$, then $v_i = \tilde{a}_i' \pm \|\tilde{a}_i'\|_2 e_i$, where the sign is chosen so as to avoid cancellation.

Now, using this decomposition $Q^T A = \begin{bmatrix} R \\ 0 \end{bmatrix}$, $Q^T = H_n \ldots H_1$,

$$x = \arg\min\|Ax - b\|_2 = \arg\min\|Q^T Ax - Q^T b\|_2 = \arg\min\left\|\begin{bmatrix} R \\ 0 \end{bmatrix}x - Q^T b\right\|_2;$$

let $Q^T b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$, $b_1 : n \times 1$, $b_2 : (m-n) \times 1$;

$$x = \arg\min\left\|\begin{bmatrix} R \\ 0 \end{bmatrix}x - \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}\right\|_2^2 = \arg\min(\|Rx - b_1\|_2^2 + \|b_2\|_2^2) = \arg\min\|Rx - b_1\|_2 = R^{-1}b_1$$

c)

	Computations	Accuracy
Normal equations	$A^T A$ - $\approx mn^2$ flops; Cholesky factorization - $\approx \frac{n^3}{3}$ flops Triangular systems - $O(n^2)$ $\approx mn^2 + \frac{n^3}{3}$ flops	relative error in x is proportional to $(cond(A))^2$
Householder transformations	$\approx 2mn^2 + \frac{2}{3}n^3$ flops	relative error in x is proportional to $cond(A) + \|r\|_2 (cond(A))^2$

For nearly square problems, $m \approx n$, the two methods require about the same amounts of work, but for $m >> n$ normal equations method is about 2 times cheaper than Householder method. On the other hand, the Householder method is more accurate.

2 a) Via elementary symbolic calculations we obtain $y(t) = e^{\lambda t}$. As $t \to \infty$, $y(t) \to +0$.

b) $y_{k+1}(1 - \frac{\lambda h}{2}) = y_k(1 + \frac{\lambda h}{2})$, $\quad y_{k+1} = \frac{1 + \lambda h / 2}{1 - \lambda h / 2} y_k$, $\quad y_k = \left(\frac{1 + \lambda h / 2}{1 - \lambda h / 2}\right)^k y_0$

$y_k \to 0 \quad \Leftrightarrow \quad \left|\frac{1 + \lambda h / 2}{1 - \lambda h / 2}\right| < 1, \quad (2 + \lambda h)^2 < (2 - \lambda h)^2, \quad 4\lambda h < -4\lambda h, \quad \lambda h < 0,$

and since we assume $h > 0$, this is true for all negative λ.

c) We have
$$y_{k+1} = y_k + \frac{f(t_k, y_k) + f(t_{k+1}, y_{k+1})}{2} h \qquad (1)$$
For the true solution $y(t)$ we can write
$$y(t_{k+1}) = y(t_k) + \frac{y'(t_k) + y'(t_{k+1})}{2} h + \xi, \qquad (2)$$
where ξ is some unknown term. From Taylor expansion we have two formulas:

$$\frac{f(x+h) + f(x-h)}{2} = f(x) + \frac{\theta h^2}{2}, \qquad |\theta| \le \|f''\|_C;$$
$$\frac{f(x+h) - f(x-h)}{2h} = f'(x) + \frac{\theta_1 h^2}{6}, \qquad |\theta_1| \le \|f^{(3)}\|_C.$$

Regrouping (2) and applying these formulas, we get:

$$\frac{\xi}{h} = \frac{y(t_{k+1}) - y(t_k)}{h} - \frac{y'(t_{k+1}) + y'(t_k)}{2} = y'(t_{k+1/2}) + \frac{\eta h^2}{6} - y'(t_{k+1/2}) - \frac{\eta_1 h^2}{2} = (\frac{\eta}{6} - \frac{\eta_1}{2})h^2, \text{ where}$$

$|\eta| \le \|y^{(3)}\|_C, \quad |\eta_1| \le \|y''\|_C$. Hence $|\xi| \le \frac{1}{2}(\|y''\|_C + \|y^{(3)}\|_C)h^3. \qquad (3)$

Now denoting $\delta_k = y_k - y(t_k)$ and subtracting (1)-(2), we get

$$\delta_{k+1} = \delta_k + \frac{f(t_k, y_k) - f(t_k, y(t_k))}{2} h + \frac{f(t_{k+1}, y_{k+1}) - f(t_{k+1}, y(t_{k+1}))}{2} h - \xi.$$

Then using Lipschitz-continuity of f,

$$|\delta_{k+1}| \le |\delta_k| + \frac{Lh}{2}|\delta_k| + \frac{Lh}{2}|\delta_{k+1}| + |\xi|, \qquad (1 - \frac{Lh}{2})|\delta_{k+1}| \le (1 + \frac{Lh}{2})|\delta_k| + |\xi|,$$

$$|\delta_{k+1}| \le \left|\frac{1+Lh/2}{1-Lh/2}\right| |\delta_k| + \frac{|\xi|}{1-Lh/2} \le \left|\frac{1+Lh/2}{1-Lh/2}\right| |\delta_k| + 2|\xi| \qquad \text{for} \quad h \le \frac{1}{L}.$$

To simplify this, consider the general situation: $x_{k+1} \le a x_k + b$. Applying this inequality recursively, we get $x_k \le a^k x_0 + b(a^{k-1} + a^{k-2} + \ldots + 1) = a^k x_0 + b\dfrac{a^k - 1}{a-1}$. The first term in our case disappears, because $y(0) = y_0 = 1$; therefore $|\delta_k| \le 2|\xi| \dfrac{\left|\frac{1+Lh/2}{1-Lh/2}\right|^k - 1}{\left|\frac{1+Lh/2}{1-Lh/2}\right| - 1}$.

Now since $k = t_k / h$, $\left|\dfrac{1+Lh/2}{1-Lh/2}\right|^k \xrightarrow[h\to 0]{} \dfrac{e^{\frac{Lt_k}{2}}}{e^{-\frac{Lt_k}{2}}} = e^{Lt_k}$, and hence

$$|\delta_k| \le 2|\xi| \frac{e^{Lt_k} + \varepsilon(h) - 1}{Lh}(1 - Lh/2) \le \frac{2|\xi|}{h} \frac{e^{Lt_k} + \varepsilon(h) - 1}{L},$$

where $\dfrac{2|\xi|}{h} \le (\|y''\|_C + \|y^{(3)}\|_C)h^2$, and $\varepsilon(h) \to 0$ as $h \to 0$. Thus $|\delta_k| \le Ch^2$ for sufficiently small h.

3 a) Let $l(x) = f(a) + (x-a)\dfrac{f(b)-f(a)}{b-a}$, then

$$I(f) = \int_a^b l(x)dx = f(a)(b-a) + \left.\frac{(x-a)^2}{2}\right|_a^b \frac{f(b)-f(a)}{b-a} = (b-a)(f(a) + \frac{f(b)-f(a)}{2}) = (b-a)\frac{f(a)+f(b)}{2}$$

---

b) Let $I_0 = \int_a^b f(x)dx$, $\quad \Delta = b - a$.

$$f(a+t) = f(a) + f'(a)t + f''(\xi)\frac{t^2}{2}, \qquad I_0 = f(a)\Delta + f'(a)\frac{\Delta^2}{2} + A\frac{\Delta^3}{6}, \qquad |A| \le \|f''\|_C$$

$$f(b-t) = f(b) - f'(b)t + f''(\xi_1)\frac{t^2}{2}, \qquad I_0 = f(b)\Delta - f'(b)\frac{\Delta^2}{2} + B\frac{\Delta^3}{6}, \qquad |B| \le \|f''\|_C .$$

Averaging the two expressions for I_0, we obtain

$$I_0 = \frac{f(a)+f(b)}{2}\Delta + \frac{A+B}{2}\frac{\Delta^3}{6} = I + C\frac{\Delta^3}{6}, \quad \text{where } C = \frac{A+B}{2}. \text{ This implies } |I_0 - I| \le \|f''\|_C \frac{\Delta^3}{6}.$$

---

c) Let $I_0^{(i)}$, $I^{(i)}$ be the true integral and our estimate correspondingly for the ith subinterval. Then

$$I_0 = \sum_{i=1}^n I_0^{(i)}, \qquad I = \sum_{i=1}^n I^{(i)}.$$

$$|I_0 - I| \le \sum_{i=1}^n |I_0^{(i)} - I^{(i)}| \underset{(b)}{\le} \sum_{i=1}^n |C_i| \frac{\Delta_i^3}{6} \le \sum_{i=1}^n \|f''\|_C \frac{1}{6}\left(\frac{\Delta}{n}\right)^3 = \|f''\|_C \frac{\Delta^3}{6n^2} = \frac{\|f''\|_C (b-a)^3}{6n^2}.$$