5.1 IVP

- $dy/dt = f(t,y),\ a \le t \le b,\ y(a) = \alpha$ is **well-posed** if it has a unique solution $y(t)$, and there are $\epsilon_0 > 0$ and $k > 0$ s.t. $\forall \epsilon \in (0,\epsilon),\ \delta_0 \in (-\epsilon, \epsilon)$, and a continuous function $\delta(t)$ satisfying $|\delta(t)| < \epsilon$, there is a unique solution to
$$\dot z = f(t,z) + \delta(t),\ a \le t \le b,\ z(a) = \alpha + \delta_0$$
satisfying $|z(t) - y(t)| < k\epsilon$ for all $t \in [a,b]$.
- When f is conti and Lipschitz in y on $D = [a,b]_t \times \mathbb{R}_y$, $\dot y = f(t,y),\ y(a) = \alpha$ is well-posed.

5.2 Euler's Method

- $w_0 = \alpha$, $w_{i+1} = w_i + hf(t_i, w_i)$.
- Err bound: if f is Lipschitz with const L on $D = [a,b]_t \times \mathbb{R}_y$ and if $|y''(t)| \le M$, then
$$|y(t_i) - w_i| \le \frac{hM}{2L}[e^{L(t_i - a)} - 1].$$
- Perturb: $u_0 = \alpha + \delta_0$, $u_{i+1} = u_i + h(t_i, u_i) + \delta_{t+1}$, then under the same hypotheses,
$$|y(t_i) - u_i| \le \frac{1}{L}\left(\frac{hM}{2} + \frac{\delta}{h}\right)[e^{L(t_i - a)} - 1] + |\delta_0| e^{L(t_i - a)}$$
where $\delta \ge \sup|\delta_i|$,

5.3 Higher-Order Taylor

- The difference method $w_0 = \alpha$, $w_{i+1} = w_i + h\phi(t_i, w_i)$ has **local trunc err**
$$\tau_{i+1}(h) = \frac{y_{i+1} - y_i}{h} - \phi(t_i, y_i), \quad y_i = y(t_i).$$
- Taylor method of order n: $w_0 = \alpha$,
$$\frac{w_{i+1} - w_i}{h} = \sum_{j=0}^{n-1} \frac{h^j}{(j+1)!} f^{(j)}(t_i, w_i).$$
Note $\dot f(t) = \partial_t f(t,y) + \partial_y f(t, y(t))\,\dot y(t)$, etc. If $y \in C^{n+1}$, then the loc trunc err is $O(h^n)$.

5.4 Runge–Kutta

- From 2nd Taylor, $T^{(2)}(t,y) \approx f(t + \frac{h}{2}, y + \frac{h}{2}f(t,y))$ gives 'midpoint method' ($O(h^2)$):
$$\frac{w_{i+1} - w_i}{h} = f\left(t_i + \frac{h}{2}, w_i + \frac{h}{2}f(t_i, w_i)\right)$$
- From 3rd Taylor, $T^{(3)}(t,y) \approx \frac{1}{2}f(t,y) + \frac{1}{2}f(t+h, y + hf(t,y))$ gives 'modified Euler method' ($O(h^2)$):
$$\frac{w_{i+1} - w_i}{h} = f(t_i, w_i) + f(t_{i+1}, w_i + hf(t_i, w_i))$$
- From 3rd Taylor, $T^{(3)}(t,y) \approx a_1 f(t,y) + a_2 f(t + \alpha_1, y + \delta_1 f(t + \alpha_2, y + \delta_2 f(t,y)))$. With proper choices, this gives 'Heun's method' ($O(h^3)$):
$$\frac{w_{i+1} - w_i}{h} = \frac{1}{4}f(t_i, w_i) + \frac{3}{4}f\left(t_i + \frac{2h}{3}, w_i + \frac{2h}{3}f\left(t_i + \frac{h}{3}, w_i + \frac{h}{3}f(t_i, w_i)\right)\right).$$
- Runge-Kutta of order 4 ($O(h^4)$): $w_0 = \alpha$,
$$\begin{aligned} k_1 &= hf(t_i, w_i), \\ k_2 &= hf(t_i + h/2, w_i + k_1/2), \\ k_3 &= hf(t_i + h/2, w_i + k_2/2), \\ k_4 &= hf(t_{i+1}, w_i + k_3), \\ w_{i+1} &= w_i + (k_1 + 2k_2 + 2k_3 + k_4)/6. \end{aligned}$$

5.6 Multistep Methods

- m-step multi method:
$$w_{i+1} = a_{m-1}w_i + \cdots + a_0 w_{i+1-m} + h\textstyle\sum_{j=0}^m b_j f(t_{i+1-m+j}, w_{i+1-j})$$
- $b_m = 0$: explicit or open;
- $b_m \ne 0$: implicit or closed.
- 4th order Adams–Bashforth (explicit)
$$\frac{w_{i+1} - w_i}{h} = \frac{55}{24}f_i - \frac{59}{24}f_{i-1} + \frac{37}{24}f_{i-2} - \frac{9}{24}f_{i-3\ge 0}$$
- 4th order Adams–Moulton (implicit)
$$\frac{\Delta w_i}{h} = \frac{9}{24}f_{i+1} + \frac{19}{24}f_i - \frac{5}{24}f_{i-1} + \frac{1}{24}f_{i-2\ge 0}$$
- Derivation of A–B: from backward diff poly,
$$f(t, y(t)) = \sum_{k=0}^{m-1}(-1)^k \binom{-s}{k}\nabla^k f(t_i, y(t_i)) + s(s+1)\cdots(s+m-1)f^{(m)}(\xi_i(s), y(\xi_i(s)))$$
where $t = t_i + sh$ so that ($\nabla p_n = p_n - p_{n-1}$)
$$\begin{aligned} &\int_{t_i}^{t_{i+1}} f(t, y(t)) \\ &= h\sum_{k=0}^{m-1}\nabla^k f(t_i, y(t_i))(-1)^k \int_0^1 \binom{-s}{k} ds \\ &+ \frac{h^{m+1}}{m!}\int_0^1 s\cdots(s+m-1)f^{(m)}(\xi_i(s), y(\xi_i(s)))ds \\ &= h\sum_{k=0}^{m-1}\nabla^k f(t_i, y(t_i))(-1)^k \int_0^1 \binom{-s}{k} ds \\ &+ h^{m+1}f^{(m)}(\mu_i, y(\mu_i))(-1)^m \int_0^1 \binom{-s}{m} ds \end{aligned}$$
so that
$$\begin{aligned} &\frac{y(t_{i+1}) - y(t_i)}{h} \\ &= \sum_{k=0}^{m-1}\left[(-1)^k \int_0^1 \binom{-s}{k} ds\right]\nabla^k f(t_i, y(t_i)) \\ &+ h^{m+1}f^{(m)}(\mu_i, y(\mu_i))(-1)^m \int_0^1 \binom{-s}{m} ds \end{aligned}$$
- Loc trunc err of multistep method:
$$\tau_{i+1} = \frac{y(t_{i+1}) - \sum_{j=0}^{m-1} a_j y(t_{i+m-1-j})}{h} - \textstyle\sum_{j=0}^m b_j f(t_{i+1-m+j}, y(t_{i+1-m+j})).$$
- m-step (m-th order) A–B: $O(h^m)$,
- $(m-1)$-step (m-th order) A–M: $O(h^m)$.
- A–B: $\frac{1}{2}$(3,-1), $\frac{1}{12}$(23,-16,5), $\frac{1}{24}$(55,-59,37,-9), $\frac{1}{720}(1901, -2774, 2616, -1274, 251)$.
- A–M: $\frac{1}{12}(5, 8, -1)$, $\frac{1}{24}(9, 19, -5, 1)$, $\frac{1}{720}(251, 646, -264, 106, -19)$.
- Predictor-Corrector method: predict w_{i+1} by A–B, correct w_{i+1} by A–M.

5.9 Highr Ord/Systems of DE

- m-th order sys of 1st order IVP:
$$\dot{\mathbf{u}}(t) = \mathbf{f}(t, \mathbf{u}(t)), \quad \mathbf{u}(a) = \boldsymbol{\alpha}.$$
If f_i are conti and Lipschitz in $\mathbf{u}$ on $D = [a,b]_t \times \mathbb{R}^m_{\mathbf{u}}$, then the IVP has a unique sol.

5.10 Stability

- A one-step diff method is consistent iff
$$\lim_{h\to 0}\max_{1\le i\le N}|\tau_i(h)| = 0.$$
- A one-step diff method is convergent iff
$$\lim_{h\to 0}\max_{1\le i\le N}|w_i - y(t_i)| = 0.$$
- A method is stable when the results depend continuously on the initial data.
- Supp a one-step diff method $w_{i+1} = w_i + h\phi(t_i, w_i, h)$ has a constant $h_0 > 0$ so that ϕ is conti and Lipschitz in w with Lipsch const L on
$$D = [a,b]_t \times \mathbb{R}_w \times [0, h_0]_h.$$
Then (a) this method is stable, (b) the method is convergent iff consistent iff $\phi(t,y,0) = f(t,y)$, (c) if there is a function τ so that $|\tau_i(h)| \le \tau(h)$ for all i and $0 \le h \le h_0$, then
$$|y(t_i) - w_i| \le \tau(h)e^{L(t_i - a)}/L.$$
- A one-step diff method is convergent iff
$$\lim_{h\to 0}\max_{1\le i\le N}|w_i - y(t_i)| = 0.$$
- A one-step diff method is consistent iff
$$\lim_{h\to 0}\max_{1\le i\le N}|\tau_i(h)| = 0,\ \lim_{h\to 0}|\alpha_i - y(t_i)| = 0.$$
- The stability of a multistep method w.r.t. round-off err is dictated by the magnitudes of the zeros of the char poly.
- Char poly: $P(\lambda) = \lambda^m - a_{m-1}\lambda^{m-1} - \cdots - a_0 = 0$ where $w_i = \alpha_i$ ($i = 0, \ldots, m-1$),
$$w_{i+1} = \textstyle\sum_{j=0}^{m-1} a_j w_{i+1-m+j} + hF(t_i, h, w_{i+1}, \ldots, w_{i+1-m})$$
If $P(\lambda) = 0 \implies |\lambda| \le 1$ and if the roots with abs value 1 are simple, we say this method satisfies the **root condition**, aka **stable**.
If 1 is the only root of char eqn with magnitude 1, **strongly stable**. O/w but satisfying the root condition, **weakly stable**. Elsewise **unstable**.
- If a multistep method is consistent, then stable iff convergent iff satisfying root condition.

5.11 Stiff DE

- The exact solution of stiff equation has term of the form e^{-ct} where c is a large positive constant, called the **transient solution**. The more important portion is called the **steady-state solution**.
- n-th derivative of e^{-ct} is $c^n e^{-ct}$ so that c^n can cause some numerical unstability.
- Test equation: $y' = \lambda y,\ y(0) = \alpha$ ($\lambda < 0$)
- Euler's method applied on test equation: we have
$$|w_i - y(t_i)| = \left|(e^{h\lambda})^i - (1 + h\lambda)^i\right||\alpha|$$
so that $|1 + h\lambda| < 1$, i.e., $h < 2/|\lambda|$ should be satisfied.
- Taylor: $\left|1 + h\lambda + \cdots + \frac{h^n\lambda^n}{n!}\right| < 1$.
- Multistep method
$$w_{i+1} = \sum_{j=0}^{m-1} a_j w_{i+1-m+j} + h\lambda\textstyle\sum_{j=0}^m b_j w_{i+1-m+j}$$
is equiv to
$$(1 - h\lambda b_m)w_{i+1} - \cdots - (a_0 + h\lambda b_0)w_{i+1-m} = 0$$
yielding the following assoc'd char poly:
$$Q(z, h\lambda) = (1 - h\lambda b_m)z^m - \cdots - (a_0 + h\lambda b_0).$$
- **Region of absolute stability**: for one-step, $R = \{w = h\lambda \in \mathbb{C} : |Q(w)| < 1\}$; for multistep, $R = \{w \in \mathbb{C} : |\beta| < 1 \text{ for all } \beta : Q(\beta, w) = 0\}$.

6.1 Gaussian Elimination

- When $a_{kk} = 0$, find minimal $k+1 \le p \le n$ and exchange k-th row with p-th row.

6.2 Pivoting Strategies

- When $a^{(k)}_{kk}$ has relatively small magnitude, errors can increase.
- Partial pivoting: choose $\max_{k\le p\le n}|a^{(k)}_{pk}|$ as a pivot elem at k-th step.
- Scaled partial pivoting: do partial pivoting after scaling each row by dividing it with $s_i = \max_{1\le j\le n}|a_{ij}|$.
- Complete pivoting: at k-th step, search all $(n+1-k)^2$ entries and select one of the largest magnitude, which yields $O(n^3)$ comparisons, while partial pivotings require $O(n^2)$ (and additional $O(n^2)$ division for scaled one).

7.1 Norms

- Matrix norm: $\|A\| \ge 0$ w/ equality iff $A = O$, $\|\alpha A\| = |\alpha|\|A\|$, $\|A + B\| \le \|A\| + \|B\|$ and $\|AB\| \le \|A\|\|B\|$.
- $\rho(A)$ is the largest abs value of eigenvalues.

- $\|A\|_2 = \sqrt{\rho(A^t A)}$ and $\rho(A) \le \|A\|$ for any induced(aka natural) norm.
- A is convergent if every entry of A^k tends to 0 as $k \to \infty$.
- A is convergent iff $\|A^n\| = 0$ for some natural norm, iff $\|A^n\| = 0$ for all natural norm, iff $\rho(A) < 1$, iff $A^n\mathbf{x} \to \mathbf{0}$ for any $\mathbf{x}$.

7.3 Jacobi/Gauss–Seidel

- Solve $A\mathbf{x} = \mathbf{b}$ by an iterative method.
- $\mathbf{x}^{(k)} = T\mathbf{x}^{(k-1)} + \mathbf{c}$.
- $A = D - L - U$, where D is diagonal part, $-L$ is strict lower triangular part and $-U$ is strict upper triangular part.
- Jacobi: $T = D^{-1}(L + U)$, $\mathbf{c} = D^{-1}\mathbf{b}$.
- Gauss-Seidel:
  $T = (D - L)^{-1}U$, $\mathbf{c} = (D - L)^{-1}\mathbf{b}$.
- Iterative method converges for any $\mathbf{x}^{(0)}$ to the unique solution to $\mathbf{x} = T\mathbf{x} + \mathbf{c}$
  $\iff \rho(T) < 1$.
- $\|x - x^{(k)}\| \le \|T\|^k \|x - x^{(0)}\|$,
- $\|x - x^{(k)}\| \le \frac{\|T\|^k}{1-\|T\|} \|x^{(1)} - x^{(0)}\|$.
- A is **diagonally dominant** iff
  $|a_{ii}| \ge \sum_{1\le j(\ne i)\le n} |a_{ij}|$. Without equality holding, A is called to be **strictly diagonally dominant**. A strictly diagonally dominant matrix is nonsingular, and in this case, (i) Gaussian elim can be done without row/column exchanges; (ii) both Jacobi and G–S works well (converge to the unique solution to $Ax = b$).
- Since $\|x^{(k)} - x\| \approx \rho(T)^k \|x^{(0)} - x\|$, we'd like to choose a method making $\rho(T) < 1$ small.
- When A has nonpositive off-diagonal entries and positive diagonal entries, then one and only one of the following holds:
  1. $0 \le \rho(T_g) < \rho(T_j) < 1$,
  2. $1 < \rho(T_j) < \rho(T_g) < 1$,
  3. $\rho(T_g) = \rho(T_j) = 0$,
  4. $\rho(T_g) = \rho(T_j) = 1$.

7.4 Relaxation Techniques

- Residual vector for $\tilde{\mathbf{x}}$: $r = \mathbf{b} - A\tilde{\mathbf{x}}$.
- Let $r_i^{(k)}$ be the res'l vec for $\tilde{\mathbf{x}}_i^{(k)}$ where
  $$\tilde{\mathbf{x}}_i^{(k)} = (x_1^{(k)}, \cdots, x_{i-1}^{(k)}, x_i^{(k-1)}, \cdots, x_n^{(k-1)})$$
  where the following holds
  $$x_i^{(k)} = x_i^{(k-1)} + r_{ii}^{(k)}/a_{ii}.$$
  Note that this choice of $x_i^{(k)}$ is making $r_{i,i+1}^{(k)} = 0$, and it is not necessarily efficient. Instead, consider the following:
  $$x_i^{(k)} = x_i^{(k-1)} + w\, r_{ii}^{(k)}/a_{ii}.$$
- $0 < w < 1$: under-relaxation;
- $w > 1$: over-relaxation (aka SOR)
- Equivalently, $\mathbf{x}^{(k)} = T_w\mathbf{x} + \mathbf{c}_w =$
$(D - wL)^{-1}[(1 - w)D + wU]\mathbf{x} + (D - wL)^{-1}w\mathbf{b}$
- How to choose w?
- If $a_{ii} \ne 0$, then $\rho(T_w) \ge |w - 1|$ so that SOR method can converge only if $0 < w < 2$.
- Converse: if A is pos def, the converse of above holds.
- If A is pos def and tridiagonal, then $\rho(T_g) = \rho(T_j)^2 < 1$ and the optimal choice is
  $$w = \frac{2}{1 + \sqrt{1 - \rho(T_j)^2}}.$$
  With this choice, $\rho(T_w) = w - 1$.

7.5 Error Bounds

- Let $\mathbf{r}$ be the residual vector for $\tilde{\mathbf{x}}$, where A is nonsingular. Then
  $$\|\mathbf{x} - \tilde{\mathbf{x}}\| \le \|\mathbf{r}\| \cdot \|A^{-1}\|$$
  and hence when $\mathbf{x} \ne \mathbf{0} \ne \mathbf{b}$,
  $$\frac{\|\mathbf{x} - \tilde{\mathbf{x}}\|}{\|\mathbf{x}\|} \le K(A)\frac{\|\mathbf{r}\|}{\|\mathbf{b}\|}$$
  where $K(A) = \|A\|\,\|A^{-1}\|$ is condition #. When ill-conditioned ($K(A) \gg 1$), making accuracy decisions based on $\|\mathbf{r}\|$ makes no sense.
- Perturb: suppose A is nonsingular and $\|\delta A\| < 1/\|A^{-1}\|$, then the solution $\tilde{\mathbf{x}}$ to $(A + \delta A)\tilde{\mathbf{x}} = \mathbf{b} + \delta\mathbf{b}$ approximates $\mathbf{x} : A\mathbf{x} = \mathbf{b}$ where
  $$\frac{\|\mathbf{x} - \tilde{\mathbf{x}}\|}{\|\mathbf{x}\|} \le \frac{K(A)\|A\|}{\|A\| - K(A)\|\delta A\|}\left(\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|} + \frac{\|\delta A\|}{\|A\|}\right).$$

7.6 Conjugate Gradient Methods

- Minimize $g(\mathbf{x}) = \langle \mathbf{x}, A\mathbf{x}\rangle - 2\langle \mathbf{x}, \mathbf{b}\rangle$.
- $$t_k = \frac{\left\langle \mathbf{v}^{(k)}, \mathbf{b} - A\mathbf{x}^{(k-1)}\right\rangle}{\left\langle \mathbf{v}^{(k)}, A\mathbf{v}^{(k)}\right\rangle},$$
  $$\mathbf{x}^{(k)} = \mathbf{x}^{(k-1)} + t_k\mathbf{v}^{(k)},$$
  and choose a new search direction $\mathbf{v}^{(k+1)}$.
- Steepest descent: $\mathbf{v}^{(k+1)} = \mathbf{r}^{(k)} = \mathbf{b} - A\mathbf{x}^{(k)}$ since $\mathbf{r} = -\frac{1}{2}\nabla g(\mathbf{x})$. But converges slowly.
- A-orthog vectors $\{\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(n)}\}$: $\left\langle \mathbf{v}^{(i)}, A\mathbf{v}^{(j)}\right\rangle = C_i\delta_{ij}$. Then the procedure stops after n steps with exact solution, assuming exact arithmetics.
  Proof: show $\mathbf{r}^{(n)}$ is orthog to all $\mathbf{v}^{(k)}$.
- Conjugate direction: choosing $\mathbf{v}^{(k)}$'s so that $\left\langle \mathbf{r}^{(k)}, \mathbf{v}^{(j)}\right\rangle = 0$ for $j = 1, \ldots, k$. In summary,
  $$t_k = \frac{\left\langle \mathbf{r}^{(k-1)}, \mathbf{r}^{(k-1)}\right\rangle}{\left\langle \mathbf{v}^{(k)}, A\mathbf{v}^{(k)}\right\rangle}, \quad \mathbf{x}^{(k)} = \mathbf{x}^{(k-1)} + t_k\mathbf{v}^{(k)},$$
  $$\mathbf{r}^{(k)} = \mathbf{r}^{(k-1)} - t_kA\mathbf{v}^{(k)}, \quad s_k = \frac{\left\langle \mathbf{r}^{(k)}, \mathbf{r}^{(k)}\right\rangle}{\left\langle \mathbf{r}^{(k-1)}, \mathbf{r}^{(k-1)}\right\rangle},$$
  $$\mathbf{v}^{(k+1)} = \mathbf{r}^{(k)} + s_k\mathbf{v}^{(k)}.$$
- Convergence rate of steepest descent:
$$g(\mathbf{x}^{(k)}) - g(\mathbf{x}^*) \le \left(\frac{K(A) - 1}{K(A) + 1}\right)^{2k}(g(\mathbf{x}^{(0)}) - g(\mathbf{x}^*)).$$
- Convergence rate of conjugate gradient ($k \le n$):
$$g(\mathbf{x}^{(k)}) - g(\mathbf{x}^*) \le 4\left(\frac{\sqrt{K(A)} - 1}{\sqrt{K(A)} + 1}\right)^{2k}(g(\mathbf{x}^{(0)}) - g(\mathbf{x}^*))$$
- Preconditioning: to increase $K(A)$.
  $$\tilde{A} = C^{-1}A(C^{-1})^t, \quad \tilde{A}(C^t\mathbf{x}) = C^{-1}\mathbf{b}.$$
- One choice: $C = \mathrm{diag}(a_{11}, \ldots, a_{nn})$.
- When A is pos def, Cholesky decomp: $A = LL^t$, let $C = L$, then $\tilde{A} = I$.

8.1 Discrete Least □'s Approx

- Minimize $E = \sum_{i=1}^m (y_i - (a_1x_i + a_0))^2$ yields ($\partial E/\partial a_j = 0$)
  $$a_0 = \frac{(\sum x_i^2)(\sum y_i) - (\sum x_iy_i)(\sum x_i)}{m(\sum x_i^2) - (\sum x_i)^2},$$
  $$a_1 = \frac{m\sum x_iy_i - (\sum x_i)(\sum y_i)}{m(\sum x_i^2) - (\sum x_iy_i)^2}$$

8.2 Orthog Poly & LSA

- Approx $f \in C[a, b]$ by $P_n(x) = \sum_0^n a_jx^j$.
  $$\sum_{k=0}^n a_k \int_a^b x^{j+k}\,dx = \int_a^b x^jf(x)\,dx$$
  for $j = 0, \ldots, n$, i.e., with $\mathbf{a} = (a_0, \ldots, a_n)^t$ and $\mathbf{b} = \left(\int_a^b x^jf(x)\,dx\right)_{j=0}^n$, $H\mathbf{a} = \mathbf{b}$ where
  $$H_{jk} = \int_a^b x^{j+k}\,dx = \frac{b^{j+k+1} - a^{j+k+1}}{j + k + 1},$$
  $0 \le j, k \le n$ is ill-conditioned.
- More efficient way: using orthog poly's ϕ_j ($j = 0, \ldots, n$), $P = \sum a_j\phi_j$ is the least square solution where
  $$a_j = \frac{\langle\phi_j, f\rangle}{\langle\phi_j, \phi_j\rangle}, \quad \langle f, g\rangle = \int_a^b wfg\,dx.$$
- Recurrence formula for the orthog poly: $\phi_0(x) = 1$, $\phi_1(x) = x - B_1$, $\phi_{k(\ge 2)}(x) = (x - B_k)\phi_{k-1}(x) - C_k\phi_{k-1}(x)$,
  $$B_k = \frac{\langle x\phi_j, \phi_j\rangle}{\langle\phi_j, \phi_j\rangle}, \quad C_k = \frac{\langle x\phi_{k-1}, \phi_{k-2}\rangle}{\langle\phi_{k-2}, \phi_{k-2}\rangle}.$$

8.3 Chebyshev Poly's

- $T_n(x) = \cos(n\arccos(x))$, orthogonal with the weight $w(x) = (1 - x^2)^{-1/2}$, $\langle T_n, T_n\rangle = \pi/2$.
- T_n has n simple zeros in $[-1, 1]$ at $\bar{x}_k = \cos(\frac{2k-1}{2n}\pi)$, and its absolute extrema at $\bar{x}'_k = \cos(k\pi/n)$ with $T_n(\bar{x}'_k) = (-1)^k$.
- Monic Chebyshev: $\tilde{T}_0 = 1$, $\tilde{T}_n = T_n/2^{n-1}$.
  $$\tilde{T}_2 = x\tilde{T}_1 - \tfrac{1}{2}\tilde{T}_0, \quad \tilde{T}_{n+1} = x\tilde{T}_n - \tfrac{1}{4}\tilde{T}_{n-1}.$$
- $\tilde{T}_n$ has minimal absolute maximum value among monic poly's of deg n on $[-1, 1]$: for monic P_n of degree n,
  $$2^{1-n} = \max_{[-1,1]} |\tilde{T}_n(x)| \le \max_{[-1,1]} |P_n(x)|.$$
- By letting x_i to be $(i + 1)$-th zero of T_{n+1}, (upp bd of) Lagrange interpolation error is minimized (on $[-1, 1]$):
  $$f(x) - P(x) = \tfrac{f^{(n+1)}(\xi(x))}{(n+1)!}(x - x_0)\cdots(x - x_n),$$
  $$\max_{[-1,1]} |f(x) - P(x)| \le \frac{\max_{[-1,1]} |f^{(n+1)}(x)|}{2^n(n + 1)}$$
- Approx P_n by $(n - 1)$ deg poly P_{n-1}:
  $$\max_{[-1,1]} \left|\tfrac{1}{a_n}(P_n(x) - P_{n-1}(x))\right| \ge \tfrac{1}{2^{n-1}}$$
  so that letting $(P_n - P_{n-1})/a_n = \tilde{T}_n$ we achieve the minimum:
  $$\max_{[-1,1]} |P_n(x) - P_{n-1}(x)| = \frac{|a_n|}{2^{n-1}}$$
  when $P_{n-1} = P_n - a_n\tilde{T}_n$.

9.1 Eigenvalues

- Geršgorin Circle theorem: let A be an $n\times n$ matrix and R_i be the circle in the complex plane with center a_{ii} and radius $\sum_{j=1,j\ne i}^n |a_{ij}|$. Then the eigenvalues of A are contained within the union of these circles, and each connected component of the union of circles contains exactly k eigenvalues where k is the # of circle merged to form the component.

9.2 Power Method

- $|\lambda_1| \ge \cdots \ge |\lambda_n|$: eigenval's of A, $\|\mathbf{x}^{(0)}\|_\infty = 1$, $\mathbf{x}^{(0)} = \sum \beta_k\mathbf{v}^{(k)}$, $\mathbf{v}^{(k)}$: unit eig'vec's corr to λ_k.
  $\mathbf{y}^{(m)} = A\mathbf{x}^{(m-1)}$,
  $$\mu^{(m)} = y_{p_{m-1}}^{(m)} = \lambda_1\left[\frac{\sum_{j=1}^n(\frac{\lambda_j}{\lambda_1})^m\beta_j\mathbf{v}_{p_{m-1}}^{(j)}}{\sum_{j=1}^n(\frac{\lambda_j}{\lambda_1})^{m-1}\beta_j\mathbf{v}_{p_{m-1}}^{(j)}}\right],$$
  $$\mathbf{x}^{(m)} = \frac{\mathbf{y}^{(m)}}{y_{p_m}^{(m)}} = \frac{A^m\mathbf{x}^{(0)}}{\|A^m\mathbf{x}^{(0)}\|_\infty}.$$
  where $|y_{p_j}^{(j)}| = \|\mathbf{y}^{(j)}\|_\infty, p_j$ min'l; $\mu^{(m)} \to \lambda_1$ and $\mathbf{x}^{(m)} \to \mathbf{v}^{(1)}$, provided by $\beta_1 \ne 0$.
- Deflation methods: matrix B with same eig'val's with A except the dominant eig'val replaced with 0.
- When λ_1 has multiplicity 1, $B = A - \lambda_1\mathbf{v}^{(1)}\mathbf{x}^t$ (with $\mathbf{x}^t\mathbf{v}^{(1)} = 1$) has eig'val's $0, \lambda_2, \ldots, \lambda_n$ with assoc eig'vec's $\mathbf{v}^{(1)}, \mathbf{w}^{(2)}, \ldots, \mathbf{w}^{(n)}$ where
  $$\mathbf{v}^{(i)} = (\lambda_i - \lambda_1)\mathbf{w}^{(i)} + \lambda_1(\mathbf{x}^t\mathbf{w}^{(i)})\mathbf{v}^{(1)}.$$
- Wielandt deflation: $\mathbf{x} = (a_{i1}, \ldots, a_{in})^t/(\lambda_1v_i^{(1)})$ provided by $v_i^{(1)} = (\mathbf{v}^{(1)})_i \ne 0$. With this, i-th row of B is a zero vector. Therefore, $B\mathbf{w} = \lambda\mathbf{w}$ implies i-th entry of $\mathbf{w}^{(j\ge 2)}$ is 0.
- After Wielandt, B' obtained from B removing i-th row and column has $\lambda_2, \ldots, \lambda_n$.
- Eigenvec for B from B': insert 0 between $(i - 1)$-th and i-th entry.