The purpose of this lab is to study least squares approximation, both discrete and continuous, using the statistical error function

\[ \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \]

(Erf[x] in Mathematica). Generate 11 data points by taking \( t_i = (i - 1)/10 \) and \( b_i = \text{erf}(t_i) \), \( i = 1, \ldots, 11 \). Make a table showing all the results from the following four parts.

(5) 1. Fit the data in a least-squares sense with polynomials of degrees from 1 to 10. Compare the fitted polynomial with \( \text{erf}(t) \) for 100 values of \( t \) between the data points, and see how the maximum (absolute) error depends on \( n \), the number of coefficients in the polynomial.

(5) 2. Since \( \text{erf}(t) \) is an odd function of \( t \), that is, \( \text{erf}(t) = -\text{erf}(-t) \), it is reasonable to fit the same data by a linear combination of odd powers of \( t \),

\[ \text{erf}(t) \approx c_1 t + c_2 t^3 + \ldots + c_n t^{2n-1}. \]

Again, see how the error between data points depends on \( n \). Since \( t \) varies over \([0, 1]\) in this problem, it is not necessary to consider using other basis polynomials.

(5) 3. Polynomials are not particularly good approximants for \( \text{erf}(t) \) because they are unbounded for large \( t \), whereas \( \text{erf}(t) \) approaches 1 for large \( t \). So, using the same data points, fit a model of the form

\[ \text{erf}(t) \approx c_1 + e^{-t^2} (c_2 + c_3 z + c_4 z^2 + c_5 z^3) \]

where \( z = 1/(1 + t) \). How does the error between data points compare with the polynomial models?

---

**Extra Credit**

(5) 4. Fit \( \text{erf}(t) \), \( 0 \leq t \leq 1 \), with a model of the form

\[ \text{erf}(t) \approx \sum_{i=0}^{n} c_i T_i(2t - 1) \]

where the \( T_i \) are Chebyshev polynomials and \( \sum c_i T_i \) is the best continuous least squares approximation with respect to the inner product

\[ \langle f, g \rangle = \int_{-1}^{1} \frac{f(x) g(x)}{\sqrt{1 - x^2}} dx. \]

(Translate \( \text{erf}(t) \), \( 0 \leq t \leq 1 \), to the interval \([-1, 1]\], erf\((x + 1)/2\), \(-1 \leq x \leq 1\), so that the Chebyshev polynomials apply.) Compare the error here to that of the previous models. To get NIntegrate to work with an improper integral, make the change of variable \( x = \cos \theta \), and use the option PrecisionGoal \( \rightarrow 9 \) to speed up the integration.

The relevant Mathematica functions are Fit, Projection, NIntegrate, and Normalize.
1. (Problem 5.27, page 209 in Shampine, Allen, and Pruess.) This exercise is representative of a great many computations arising in the use of the classical separation of variables technique for solving field problems. Typically, one must compute many roots of nonlinear equations and integrals. The temperature distribution in a cylinder of radius $a$ and height $b$ with the bottom ($z = 0$) held at a temperature zero, the top ($z = b$) at a temperature $f(r)$, and the side dissipating heat according to Newton’s law of cooling, can be represented by a series. If the thermal conductivity of the cylinder is $k$ and the thalpance is $\epsilon$, then the temperature $\phi(r, z)$ at radius $r$ and height $z$ is

$$\phi(r, z) = \sum_{n=1}^{\infty} A_n \frac{\sinh(q_n z)}{\sinh(q_n b)} J_0(q_n r).$$

The numbers $q_n$ are the positive roots of the equation

$$\frac{k}{ea} q_n a J_1(q_n a) - J_0(q_n a) = 0,$$

where the functions $J_0(x)$ and $J_1(x)$ are Bessel functions of the first kind of orders zero and one, respectively. The coefficients $A_n$ are given by

$$A_n = \frac{2}{a^2 \left[ 1 + \left( \frac{kq_n}{\epsilon} \right)^2 \right] J_1^2(q_n a)} \int_0^a r f(r) J_0(q_n r) dr.$$

The roots $q_n$ depend only on the geometry and the material. Once they have been computed, one can consider virtually any temperature distribution $f(r)$ by computing the quantities $A_n$. For $k/ea = 2$, compute the roots $q_n a$ for $n = 1, 2, 3$. (For your information, $q_3 a = 7.08638084796766$.) Then for $a = 1$, compute $A_1, A_2, A_3$ for $f(r) = \exp(-r) - \exp(-1)$. Now for $b = 3$, estimate the temperature $\phi(r, z)$ halfway up the outside of the cylinder ($r = 1, z = 1.5$).

Suggestion: use the root finding routine ZEROIN (cs3410 archive), integration routine ADAPT (cs3410 archive), and DBESJ (kahaner archive) for the Bessel functions. The entire exercise can also be done in Mathematica.
THIS CASE FORWARD RECURSION IS STABLE AND VALUES FROM THE
ASYMPTOTIC EXPANSION FOR X TO INFINITY START THE RECURSION
WHEN IT IS EFFICIENT TO DO SO. LEADING TERMS OF THE SERIES
AND UNIFORM EXPANSION ARE TESTED FOR UNDERFLOW. IF A SEQUENCE
IS REQUESTED AND THE LAST MEMBER WOULD UNDERFLOW, THE RESULT
IS SET TO ZERO AND THE NEXT LOWER ORDER TRIED, ETC., UNTIL A
MEMBER COMES ON SCALE OR ALL MEMBERS ARE SET TO ZERO.
OVERFLOW CANNOT OCCUR.

INPUT
X - X .GE. 0.0E0
ALPHA - ORDER OF FIRST MEMBER OF THE SEQUENCE,
        ALPHA .GE. 0.0E0
N - NUMBER OF MEMBERS IN THE SEQUENCE, N .GE. 1

OUTPUT
Y - A VECTOR WHOSE FIRST N COMPONENTS CONTAIN
    VALUES FOR J/SUB(ALPHA+K-1)/(X), K=1,...,N
NZ - NUMBER OF COMPONENTS OF Y SET TO ZERO DUE TO
     UNDERFLOW,
     NZ=0 , NORMAL RETURN, COMPUTATION COMPLETED
     NZ .NE. 0, LAST NZ COMPONENTS OF Y SET TO ZERO,
         Y(K)=0.0E0, K=N-NZ+1,...,N.