Background

1. Consider the mapping from quadratic polynomials to cubic polynomials given by $p(x) \mapsto xp(x)$. With respect to the power basis $\{1, x, x^2, x^3\}$, what is the matrix associated with this mapping?

   Answer:
   $$A = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

2. Consider the mapping from functions of the form $f(x, y) = c_1 + c_2 x + c_3 y$ to values at (x_1, y_1), (x_2, y_2), and (x_3, y_3). What is the associated matrix? How would you set up a system of equations to compute the coefficient vector c associated with a vector b of function values at the three points?

   Answer: We have $Ac = b$ where
   $$A = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}.$$

3. Consider the L^2 inner product between quadratic polynomials on the interval $[-1, 1]$:
   $$\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)\,dx$$
   If we write the polynomials in terms of the power basis $\{1, x, x^2\}$, what is the matrix associated with this inner product (i.e. the matrix A such that $c_p^T A c_q = \langle p, q \rangle$ where c_p and c_q are the coefficient vectors for the two polynomials.

   Answer:
   $$A = \int_{-1}^{1} \begin{bmatrix} 1 \cdot 1 & 1 \cdot x & 1 \cdot x^2 \\ x \cdot 1 & x \cdot x & 1 \cdot x^2 \\ x^2 \cdot 1 & x^2 \cdot x & x^2 \cdot x^2 \end{bmatrix} dx = \begin{bmatrix} 2 & 0 & 2/3 \\ 0 & 2/3 & 0 \\ 2/3 & 0 & 2/5 \end{bmatrix}$$

4. Consider the weighted max norm
   $$\|x\| = \max_j w_j |x_j|$$

where $w_1, \ldots, w_n$ are positive weights. For a square matrix A, what is the operator norm associated with this vector norm?

Answer: Wrte the norm as $\|x\| = \|Wx\|_\infty$ where W is the diagonal matrix of weights. Then

$$\|A\| = \max_{\|Wx\|_\infty=1} \|WAx\|_\infty = \max_{\|y\|_\infty=1} \|WAW^{-1}y\|_\infty = \max_i \sum_j |a_{ij}|w_i/w_j.$$

5. If A is symmetric and positive definite, argue that the eigendecomposition is the same as the singular value decomposition.

   Answer: $A = Q\Lambda Q^T$ is a product of orthogonal, diagonal with positive diagonal elements, and orthogonal transpose. So it's an SVD.

6. Consider the block matrix

$$M = \begin{bmatrix} A & B \\ B^T & D \end{bmatrix}$$

   where A and D are symmetric and positive definite. Show that if

$$\lambda_{\min}(A)\lambda_{\min}(D) \geq \|B\|_2^2$$

   then the matrix M is symmetric and positive definite.

   Answer: Using $x^T Ax \geq \lambda_{\min}(A)\|x\|^2$ and norm bounds on terms involving B, we have for x and y not both zero

$$\begin{aligned}
\begin{bmatrix} x \\ y \end{bmatrix} \begin{bmatrix} A & B \\ B^T & D \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= x^T Ax + 2(x^T By) + y^T Dy \\
&\geq \lambda_{\min}(A)\|x\|^2 - 2\|x\|\|y\|\|B\| + \lambda_{\min}(D)\|y\|^2 \\
&\geq (\lambda_{\min}(A)\|x\| - \lambda_{\min}(D)\|y\|)^2 > 0
\end{aligned}$$

7. Suppose D is a diagonal matrix such that $AD = DA$. If $a_{ij} \neq 0$ for $i \neq j$, what can we say about D?

   Answer:

$$(DA)_{ij} = d_{ii}a_{ij} \text{ and } (AD)_{ij} = a_{ij}d_{jj}$$

   So if both are the same and $a_{ij} \neq 0$, then $d_{ii} = d_{jj}$.

8. Convince yourself that the product of two upper triangular matrices is itself upper triangular.

   Answer: Verify directly for the block 2-by-2 case, and the general case follows.

9. Suppose Q is a differentiable *orthogonal* matrix-valued function. Show that $\delta Q = QS$ where S is skew-symmetric, i.e. $S = -S^T$.

   Answer: Apply implicit differentiation to $Q^T Q = I$ to get

   $$(\delta Q)^T Q + Q^T(\delta Q) = 0,$$

   and note that $(\delta Q)^T Q = -(Q^T(\delta Q))^T$. Therefore, $S = Q^T(\delta Q)$ is skew symmetric.

10. Suppose $Ax = b$ and $(A + D)y = b$ where A is invertible and D is relatively small. Assuming we have a fast way to solve systems with A, give an algorithm to compute y to within an error of $O(\|D\|^2)$ in terms of two linear systems involving A and a diagonal scaling operation.

    Answer: Note that $(A+D)^{-1} = A^{-1} - A^{-1}DA^{-1} + O(\|D\|^2)$, and we can compute the first two truncated Taylor expansions for x as

    ```
    x0 = solveA(b);
    x1 = x0−solveA(d.∗x0);
    ```

11. Suppose $r = b - A\hat{x}$ is the residual associated with an approximate solution $\hat{x}$. The *maximum componentwise relative residual* is

    $$\max_i |r_i|/|b_i|.$$

    How can this be written in terms of a norm?

    Answer: This is $\|Db\|_\infty$ where D is a diagonal matrix with $1/|b_i|$ on the diagonal.

Interations in 1D

1. Consider the fixed point iteration $x_{k+1} = g(x_k)$ and assume x_* is an attractive point. Also assume $|g''(x)| < M$ everywhere. We know that

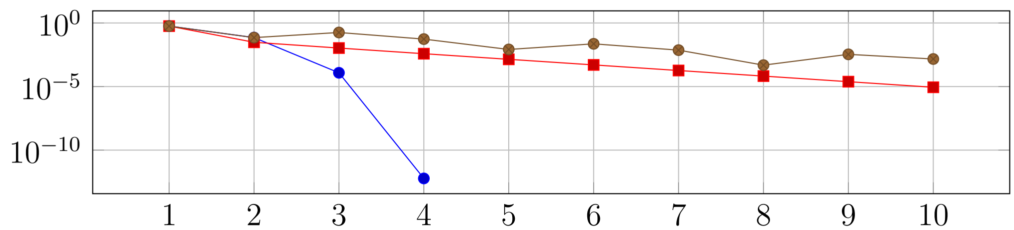

Figure 1: Convergence of Newton iteration, a fixed point iteration, and bisection for $\cos(x) = 0$.

the iteration converges to x_* from "close enough" starting points; show that a sufficient condition for convergence is

$$|x_0 - x_*| < \frac{2(1 - |g'(x_*)|)}{M}.$$

Answer: The error iteration is

$$e_{k+1} = g(x_* + e_k) - x_* = g'(x_*)e_k + \frac{1}{2}g''(\xi)e_k^2$$

Therefore

$$|e_{k+1}| \leq \left| g'(x_*) + \frac{M|e_k|}{2} \right| |e_k|,$$

and the error diminishes provided

$$|g'(x_*)| + \frac{M}{2}|e_k| < 1,$$

i.e. $|e_0| > |e_1| > \ldots > |e_k|$ under the given hypothesis.

2. What is Newton's iteration for finding $\sqrt{a}$?

   Answer: Solve $f(x) = x^2 - a = 0$:

$$x_{k+1} = x_k - \frac{x_k^2 - a}{2x_k} = \frac{x_k}{2} + \frac{a}{2x_k}.$$

3. Consider the fixed-point iteration $x_{k+1} = x_k + \cos(x_k)$. Show that for x_0 near enough to $x_* = \pi/2$, the iteration converges, and describe the convergence behavior.

Answer: The error iteration is

$$e_{k+1} = e_k - \sin(x_*)e_k + O(e_k^2) = O(e_k^2).$$

That is, we have quadratic convergence.

4. The graphs shown in Figure 1 show the convergence of Newton's iteration starting from $x_0 = 1$, the fixed point iteration $x_{k+1} = x_k + \cos(x_k)/x_k$ starting from $x_0 = 1$ and bisection starting from $[0, 2]$ to the solution of $\cos(x) = 0$. Which plot corresponds to which method? How can you tell?

   Answer: Newton's iteration converges quadratically – it goes to zero fastest (blue). Bisection gains one bit per step in general, and the error can wobble – it's the green circles. The fixed point iteration converges linearly and consistently, and is the red squares.

5. Find an example of a function with a unique zero and a starting value such that Newton's iteration does not converge.

   Answer: Something that descends to an asymptote at infinity works well; try
   $$f(x) = x\exp(-x)$$
   for $x_0 > 1$. Note for $x_k > 1$,
   $$x_{k+1} = x_k - \frac{x_k}{1 - x_k} > x_k.$$

6. Suppose f has a sign change for between $a = 1000$ and $b = 1001$. How many steps of bisection are required to obtain a *relative* error of 10^{-6}?

   Answer: A relative error of 10^{-6} corresponds to an absolute error around 10^{-3}. Since the initial interval has width 1, we start with $x_0 = 1000.5$ with error bounded by 2^{-1}. The iterate x_9 is guaranteed to have error bounded by $2^{-10} \approx 10^{-3}$.

Linear systems

1. Suppose A is square and singular, and consider $y = Ax$. show by example that a *finite* relative error in the input x can lead to an *infinite* relative error in the output y.

Answer: Consider

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad x = \begin{bmatrix} 0 & 1 \end{bmatrix}, \quad \hat{x} = \begin{bmatrix} \epsilon & 1 \end{bmatrix}.$$

2. Give a 2×2 example for which an $O(\epsilon_{\text{mach}})$ normwise relative residual corresponds to a normwise relative error near one.

   Answer: We need something ill conditioned; try

$$A = \begin{bmatrix} 1 & 0 \\ 0 & \epsilon_{\text{mach}} \end{bmatrix}, \quad x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \hat{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

3. Show that $\kappa_2(A) = 1$ iff A is a scalar multiple of an orthogonal matrix.

   Answer: If $\|A\|\|A^{-1}\| = \sigma_{\max}(A)/\sigma_{\min}(A) = 1$, then the singular values are all the same, so the SVD of A is

$$A = U(\alpha I)V^T = \alpha U V^T$$

   i.e. A is a multiple of an orthogonal matrix.

4. Suppose M is the elementary transformation matrix

$$M = \begin{bmatrix} 1 & 0 \\ m & I \end{bmatrix}.$$

   What is M^{-1}?

   Answer:

$$M = \begin{bmatrix} 1 & 0 \\ -m & I \end{bmatrix}.$$

5. Compute the Cholesky factorization of the matrix

$$A = \begin{bmatrix} 4 & 2 \\ 2 & 9 \end{bmatrix}$$

   Answer:

$$R = \begin{bmatrix} 2 & 1 \\ 0 & 2\sqrt{2} \end{bmatrix}$$

6. Consider the matrix
$$\begin{bmatrix} D & u \\ u^T & \alpha \end{bmatrix}$$
where D is diagonal with positive diagonal elements larger than the corresponding entries of u. For what range of α must u be positive definite?

   Answer: Cholesky completes if
$$\alpha > \sum_j u_j^2/d_j.$$

7. If A is symmetric and positive definite with Cholesky factor R, show that $\kappa_2(A) = \kappa_2(R)^2$ (note: use the SVD).

   Answer: Let $R = U\Sigma V^T$. Then $A = R^T R = V\Sigma^2 V^T$ gives the SVD of A.

8. If $\hat{A} = LU = A + E$, show that iterative refinement with the computed LU factors satisfies
$$\|e_{k+1}\| \leq \left(\kappa(\hat{A}) \frac{\|E\|}{\|\hat{A}\|} \right) \|e_k\|$$

   Answer: The error iteration is
$$e_{k+1} = e_k - \hat{A}^{-1} A e_k = \hat{A}^{-1} E e_k.$$
   Norm bounds give
$$\|e_{k+1}\| = \|\hat{A}^{-1} E e_k\| \leq \|\hat{A}^{-1}\| \|E\| \|e_k\| \leq \kappa(\hat{A}) \frac{\|E\|}{\|A\|} \|e_k\|.$$

Least squares problems

1. Suppose M is symmetric and positive definite, so that $\|x\|_M = \sqrt{x^T M x}$ is a norm. Write the normal equations for minimizing $\|Ax - b\|_M^2$.

   Answer:
$$A^T M (Ax - b) = 0.$$

2. Suppose $A \in \mathbb{R}^{n \times 1}$ is a vector of all ones. Show that $A^\dagger b$ is the sample mean of the entries of b.

   Answer: Write the normal equations
$$A^T A x = A^T b$$
   and note that $A^T A = n$ and $A^T b = \sum_j b_j$. Thus
$$x = \sum_j b_j / n$$
   is the sample mean.

3. Suppose $A = QR$ is an economy QR decomposition. Why is $\kappa(A) = \kappa(R)$?

   Answer: Write the SVD of R as $R = U\Sigma V^T$. Then the SVD of A is $(QU)\Sigma V^T$. The two matrices have the same singular values, hence the same condition number.

4. Suppose we have economy QR decompositions for $A_1 \in \mathbb{R}^{m_1 \times n}$ and $A_2 \in \mathbb{R}^{m_2 \times n}$, i.e.
$$A_1 = Q_1 R_1, \quad A_2 = Q_2 R_2$$
   Show that we can compute the QR decomposition of A_1 and A_2 stacked as
$$\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = QR, \quad Q = \frac{1}{\sqrt{2}} \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \tilde{Q}$$
   where
$$\tilde{Q} R = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}$$
   is an economy QR decomposition.

   Answer:
$$\begin{aligned} \tilde{Q}^T \tilde{Q} &= \left(\frac{1}{\sqrt{2}} \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \tilde{Q} \right)^T \left(\frac{1}{\sqrt{2}} \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} \tilde{Q} \right) \\ &= \frac{1}{2} \tilde{Q}^T (Q_1^T Q_1 + Q_2^T Q_2) \tilde{Q} \\ &= \frac{1}{2} \tilde{Q}^T (2I) \tilde{Q} \\ &= \tilde{Q}^T \tilde{Q} = I \end{aligned}$$

5. Give an example of $A \in \mathbb{R}^{2\times 1}$ and $b \in \mathbb{R}^2$ such that a small relative change to b results in a large relative change to the solution of the least squares problem. What is the condition number of A?

   Answer: Consider

$$A = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad b = \begin{bmatrix} \epsilon \\ 1 \end{bmatrix}, \quad \hat{b} = \begin{bmatrix} 2\epsilon \\ 1 \end{bmatrix},$$

   The condition number of A is one, but changing from b to $\hat{b}$ leads to a 100% change in the least squares solution (from ϵ to 2ϵ).

6. Write the normal equations for a Tikhonov-regularized least squares problem.

   Answer:

$$(A^T A + \lambda I)x = A^T b$$

7. Show that $\Pi = AA^\dagger$ is a projection ($\Pi^2 = \Pi$) and that Πb is the closest point to b in the range of A.

   Answer: Note that $A^\dagger A = (A^T A)^{-1} A^T A = I$, so

$$\Pi^2 = AA^\dagger AA^\dagger = AA^\dagger = \Pi.$$

   Also, $\Pi b - b = r$ is orthogonal to the span of A.

8. Using the normal equations approach, find the coefficients α and β that minimize

$$\phi(\alpha, \beta) = \int_{-1}^{1} (\alpha + \beta x - f(x))^2 \, dx$$

   Answer:

$$\left(\int_{-1}^{1} \begin{bmatrix} 1 & x \end{bmatrix}^T \begin{bmatrix} 1 & x \end{bmatrix} \, dx \right) \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \int_{-1}^{1} \begin{bmatrix} 1 & x \end{bmatrix}^T f(x) \, dx$$

   This gives

$$\begin{bmatrix} 2 & 0 \\ 0 & 2/3 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \int_{-1}^{1} f(x) \, dx \\ \int_{-1}^{1} x f(x) \, dx \end{bmatrix}$$

Eigenvalues

1. The *spectral radius* of a matrix A is the maximum modulus of any of its eigenvalues. Show that $\rho(A) \leq \|A\|$ for any operator norm.

   Answer: There exists some unit eigenvector associated wth the largest modulus eigenvalue (λ) such that

   $$\rho(A) = \|\lambda v\| = \|Av\| \leq \|A\|.$$

2. Suppose $A \in \mathbb{R}^{n\times n}$ is a symmetric matrix an $V \in \mathbb{R}^{n\times n}$ is invertible. Show that A is positive definite, negative definite, or indefinite iff $V^T AV$ is positive definite, negative definite, or indefinite.

   Answer: Define $y = V^{-1}x$ and notice that $y \neq 0$ iff $x \neq 0$. Then $y^T(V^T AV)y = x^T Ax$. Therefore, $x^T Ax$ is positive for all nonzero x iff $y^T(V^T AV)y$ is positive for all nonzero y; and similarly, if $x^T Ax$ is ever negative, so is the corresponding $y^T(V^T AV)y$.

3. Write a MATLAB fragment to take `numiter` steps of shift-invert iteration with a given shift. You should make sure that the cost per iteration is $O(n^2)$, not $O(n^3)$.

   Answer:

   ```
   [L,U,P] = lu(A-sigma*I);
   for k = 1:numiter
     v = U\(L\(P*v));
     v = v/norm(v);
   end
   ```

4. Suppose T is a block upper-triangular matrix with diagonal blocks in $\mathbb{R}^{1\times 1}$ or $\mathbb{R}^{2\times 2}$. Show that the eigenvalues of T are the diagonal values in the 1×1 blocks together with the eigenvalue pairs from the 2×2 blocks.

   Answer: A block upper triangular matrix is singular iff a diagonal block is singular. In this case, this means $T - \lambda I$ is singular iff λ is equal to one of the 1-by-1 blocks or is an eigenvalue of a 2-by-2 block.

5. If $AU = UT$ is a complex Schur form, argue that $A^{-1}U = UT^{-1}$ is the corresponding complex Schur form for A^{-1}.

Answer: Since $A = UTU^*$, we have

$$A^{-1} = U^{-*}T^{-1}U^{-1} = UT^{-1}U^*$$

and the inverse of an upper triangular matrix is upper triangular.

6. Suppose Q_k is the kth step of a subspace iteration, and Q_* is an orthonormal basis for the subspace to which the iteration is converging. Let θ be the biggest angle between a vector in the range of Q_* and the best approximation by a vector in the range of Q_k, and show that $\cos(\theta)$ is the smallest singular value of $Q_k^T Q_*$.

   Answer: Let $Q_k u$ and $Q_* v$ be vectors in the two subspaces. The cosine of the angle between the two is

   $$\cos(\theta_{uv}) = \frac{u^T Q_k^T Q_* v}{\|u\|\|v\|}.$$

   Note that if $Q_k^T Q_* = U\Sigma V^T$, we can write

   $$\cos(\theta_{uv}) = \frac{\tilde{u}^T \Sigma \tilde{v}}{\|\tilde{u}\|\|\tilde{v}\|}$$

   where $\tilde{u} = Uu$ and $\tilde{v} = Vv$. For a given $\tilde{v}$, the value of $\tilde{u}$ that maximizes $\cos(\theta_{uv})$ is $\tilde{v} = \tilde{u}$, yielding weighted average of the singular values. Setting $\tilde{u} = e_n$ gives the smallest singular value, which is thus the required cosine.

7. Show that the power method for the Cayley transform matrix $(\sigma I + A)(\sigma I - A)^{-1}$ for $\sigma > 0$ will first converge to an eigenvalue of A with positive real part, assuming such an eigenvalue exists and the iteration converges at all.

   Answer: By the spectral mapping theorem, the eigenvalues of the transformed matrix are

   $$\mu = \frac{\sigma + \lambda}{\sigma - \lambda},$$

   which is less than one in magnitude for eigenvalues λ of the matrix A that are in the left half plane, and greater than one in magnitude for eigenvalues of A in the right half plane.

8. In control theory, one often wants to plot a *transfer function*

$$h(s) = c^T(A - sI)^{-1}b.$$

The transfer function can be computed in $O(n^2)$ time using a Hessenberg reduction on A. Describe how.

Answer: Let $A = QHQ^T$. Then

$$h(s) = (c^TQ)(H - sI)^{-1}(Q^Tb),$$

and we can pre-compute Q^Tb and Q^Tc independent of s, and solve the remaining Hessenberg system $(H - sI)w = Q^Tb$ in $O(n^2)$ time.

Stationary iterations

1. Consider using Richardson iteration to solve the problem $(I - K)x = b$ where $\|K\| < 1$ (i.e. $M = I$). If $x_0 = 0$, show that x_k corresponds to taking k terms in a truncated geometric series (a.k.a a Neumann series) for $(I - K)^{-1}$.

   Answer: Richardson iteration uses the splitting $M = I$ and $K = K$, i.e.

   $$x_{k+1} = Kx_k + b$$

   Thus, $x_1 = b$, $x_2 = Kb + b$, $x_3 = K^2b + Kb + b$, and in general

   $$x_k = \sum_{j=0}^{k-1} K^j b,$$

   which is a partial sum of the Neumann series.

2. If A is strictly *column* diagonally dominant, Jacobi iteration still converges. Why?

   Answer:

3. Show that if A is symmetric and positive definite and x_* is a minimizer for the energy

   $$\phi(x) = \frac{1}{2}x^TAx - x^Tb$$

then

$$\phi(x) - \phi(x_*) = \frac{1}{2}(x - x_*)^T A(x - x_*).$$

Answer: By stationarity, of x_*, $Ax_* = b$. Therefore

$$\phi(x) - \phi(x_*) = \frac{1}{2}(x^T Ax - x_*^T Ax_*) - (x - x_*)^T b \quad = \frac{1}{2}(x^T Ax - x_*^T Ax_*) - x^T Ax_* + x_*^T Ax_* =$$

4. The largest eigenvalue of the tridiagonal matrix $T \in \mathbb{R}^{n\times n}$ is $2-O(n^{-2})$. Argue that the iteration matrix for Jacobi iteration therefore has spectral radius $\rho(R) = 1 - O(n^{-2})$, and therefore

   $$\log \rho(R) = -O(n^{-2})$$

   Using this fact, argue that it takes $O(n^2)$ Jacobi iterations to reduce the error by a constant factor for this problem.

   Answer: The matrix is positive definite, so the spectral radius is the largest eigenvalue. The diagonal part is just 2, so the Jacobi iteration matrix is $T/2$, which has largest eigenvalue $1 - O(n^{-2})$. The error behaves like

   $$\|e_m\| \leq \|T^m\|\|e_0\| = (1 - O(n^{-2}))^m\|e_0\|;$$

   taking logs, we have

   $$\log \|e_m\| \leq \log \|e_0\| - mO(n^{-2})$$

   To reduce the log error by a constant factor therefore requires $m = O(n^2)$ iterations.

Krylov subspace methods

1. Suppose A is symmetric positive definite and $\phi(x) = x^T Ax/2 - x^T b$. Show that over all approximations of the form $\hat{x} = Uz$, the one that minimizes ϕ satisfies $(U^T AU)z = U^T b$.

   Answer: Consider

   $$\phi(Uz) = \frac{1}{2}z^T(U^T AU)z - z^T(U^T b),$$

   The gradient with respect to z is zero precisely at

   $$(U^T AU)z = U^T b.$$

2. Suppose A is SPD and ϕ is defined as in the previous question. If $\hat{x} = Uz$ minimizes the energy of $\phi(\hat{x})$, show that $\|\hat{x} - x_*\|_A^2$ is also minimal.

   Answer: Observe from an earlier exercise that

   $$\frac{1}{2}\|x - x_*\|_A^2 = \phi(x) - \phi(x_*)$$

   Therefore, minimizing $\phi(x)$ (or $\phi(x) - \phi(x_*)$) over the subspace is equivalent to minimizing the energy error.

3. Suppose A is nonsingular and has k distinct eigenvalues. Argue that $\mathcal{K}_k(A, b)$ contains $A^{-1}b$.

   Answer: The conditions $p(\lambda) = \lambda^{-1}$ for all k eigenvalues can be satisfied for a polynomial p with degree at least $k - 1$. Therefore, $A^{-1}b = p(A)b \in \mathcal{K}_k(A, b)$.

4. Argue that the residual after k steps of GMRES with a Jacobi preconditioner is no larger than the residual after k steps of Jacobi iteration.

   Answer: GMRES minimizes the residual over the space spanned by k steps of the Jacobi iteration, and the minimum residual over the space must be smaller than the residual for the last vector in a basis for that space.

5. If A is symmetric, the largest eigenvalue is the maximum value of the Rayleigh quotient $\rho_A(x)$. Show that computing the largest eigenvalue of $\rho_T(z)$ where $T = Q^T AQ$ is equivalent to maximizing $\rho_A(x)$ over x s.t. $x = Qz$. The largest eigenvalue of T is always less than or equal to the largest eigenvalue of A; why?

   Answer: Assume Q has orthonormal columns; then

   $$\rho_T(z) = \frac{z^T Qz}{z^T z} = \frac{z^T Q^T AQz}{z^T Q^T Qz} = \rho_A(Qz).$$

   Maximizing $\rho_T(z)$ over z gives the largest eigenvalue of T; but $\rho_T(z) = \rho_A(Qz) \leq \lambda_{\max}(A)$. Thus the largest eigenvalue of T is bounded by the largest eigenvalue of A.

From linear to nonlinear

1. Write a MATLAB code to estimate α and x such that $y = \alpha x$ is tangent to $y = \cos(x)$ near $x_0 = n\pi$ for $n > 0$. I recommend writing two equations (matching function values and matching derivatives) in two unknowns (the intersection x and α) and applying Newton. What is a good initial guess?

   Answer: Both the values and the derivatives of $\cos(x)$ and αx should agree at the tangent point:

   $$F(\alpha, x) = \begin{bmatrix} \alpha x - \cos(x) \alpha + \sin(x) \end{bmatrix} = 0.$$

   The Jacobian is
   $$J = \begin{bmatrix} x & \alpha + \sin(x) \\ 1 & \cos(x) \end{bmatrix}$$

   A good initial guess is $x_0 = n\pi$ and $\alpha = (-1)^n/(n\pi)$.

   The code would be

   ```
   function [alpha, x] = revtangent(n)

     x = n*pi;
     alpha = (-1)^n/x;
     u = [alpha; x];
     for k = 1:20
       F = [u(1)*u(2)-cos(u(2)); u(1) + sin(u(2))];
       J = [u(2), u(1) + sin(u(2)); 1, cos(u(1))];
       u = u-J\F;
       if norm(F) < 1e-10
         break;
       end
     end
     alpha = u(1);
     x = u(2);
   ```

2. Write a MATLAB code to find a critical point of $\phi(x, y) = -\exp(x^2 + y^2)(x^2 + y^2 - 2(ax + by) + c)$ using Newton's iteration.

Answer: Replace (x, y) and (a, b) with vectors, and rewrite this as

$$\begin{aligned} \phi(x) &= -s(x)\psi(x), \\ s(x) &= \exp(x^T x) \\ \psi(x) &= x^T x - 2a^T x + c \end{aligned}$$

Then differentiate the pieces:

$$\begin{aligned} \nabla s &= 2x \exp(x^T x) & H_s &= 2(I + xx^T) \exp(x^T x) \\ \nabla \psi &= 2(x - a) & H_\psi &= 2I \end{aligned}$$

Put it together to get

$$\nabla \phi = -(\nabla s)\psi - s(\nabla \psi)$$

and

$$H_\phi = -H_s \psi - (\nabla s)(\nabla \psi)^T - (\nabla \psi)(\nabla s)^T - s H_\psi.$$

After working out the calculus, the Newton iteration is pretty straightforward.

3. Write a MATLAB fragment to minimize $\sum_j \exp(r_j^2) - 1$ where $r = Ax - b$. Use a Gauss-Newton strategy (no need to bother with safeguards like line search).

   Answer:

```
for k = 1:maxiter
  F = exp(r.^2)-1;
  J = diag(2*r.*F) * A;
  p = -J\F;
  x = x + p;
  r = A*x-b;
end
```

4. Consider the fixed point iteration

$$x_{k+1} = x_k - A^{-1} F(x_k)$$

   where F has two continuous derivatives and A is some (possibly crude) approximation to the Jacobian of F at the solution x_*. Under what conditions does the iteration converge?

Answer: Expanding $F(x_* + e_k) = J_* e_k + O(\|e_k\|^2)$, the error iteration is

$$e_{k+1} = e_k - A^{-1} J_* e_k + O(\|e_k\|^2). \quad = (I - A^{-1} J_*) e_k + O(\|e_k\|^2)$$

The iteration converges if $\rho(I - A^{-1} J_*) < 1$.

5. Suppose x_* is a strong local minimum for ϕ, i.e. $\nabla\phi(x_*) = 0$ and $H_\phi(x_*)$ is positive definite. For starting points x_0 close enough to x_*, Newton with line search based on the Armijo condition behaves identically to an unguarded Newton iteration with no line search. Why?

   Answer: The idea: There is no need to cut the step when the linear model is a good enough approximation that it predicts what will happen.

6. Argue that for large enough λ, $p = -(H_\phi(x) + \lambda I)^{-1} \nabla\phi(x)$ is guaranteed to be a descent direction, assuming x is not a stationary point.

   Answer: All we need is that $\lambda > -\lambda_{\min}(H_\phi)$ in order to guarantee that the shifted matrix is positive definite.

7. Suppose $F : \mathbb{R}^n \times \mathbb{R} \mapsto \mathbb{R}^n$ (i.e. $F = F(x, s)$). If we solve $F(x(s), s) = 0$ for a given s using Newton's iteration and we are able to compute $\partial F/\partial s$ in at most $O(n^2)$ time, we can compute dx/ds in $O(n^2)$ time. How?

   Answer: Differentiate the governing equations to find

   $$\frac{\partial F}{\partial x}\frac{dx}{ds} + \frac{\partial F}{\partial s} = 0,$$

   and notice that

   $$\frac{dx}{ds} = -\left(\frac{\partial F}{\partial x}\right)^{-1} \frac{\partial F}{\partial s}$$

   involves a solve with the Jacobian that appears in Newton's iteration. If we compute a factorization at each step of Newton, therefore, we can re-use the factorization to solve this linear system in $O(n^2)$ time.

8. Describe an algorithm to minimize $\|Ax - b\|^2$ subject to $Cx = d$, where $A \in \mathbb{R}^{m \times n}$, $m > n$ and $C \in \mathbb{R}^{p \times m}$, $p < m$.

Answer: There are several ways to do this, but perhaps the simplest is to write the normal equations with Lagrange multipliers:

$$\begin{bmatrix} A^T A & C^T \\ C & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} A^T b \\ d \end{bmatrix}.$$

An alternative is the *null space* approach: write the full QR decomposition $C^T = QR$ and let $x = Qy$. Then $Cx = R_1^T y_1 = d$ is determined by the constraints, and we can compute y_2 by solving the unconstrained least squares problem of minimizing $\|AQ_2 y_2 - (b - AQ_1 y_1)\|^2$. The null space approach is fine for small problems, but we would usually prefer something like the multiplier approach when A is large and sparse.

9. Describe a fast algorithm to solve

$$Ax = b(x_n)$$

where $A \in \mathbb{R}^{n \times n}$ is a fixed matrix and $b : \mathbb{R} \to \mathbb{R}^n$ is twice differentiable. Your algorithm should cost $O(n^2)$ per step and converge quadratically.

Answer: Let $F = Ax - b(x_n)$, and note that the Jacobian is

$$J = A - b'(x_n) e_n^T,$$

which is independent of x except in the last column. The Newton update eqution is thus

$$(A - b' e_n^T) p = -(Ax - b)$$

or

$$(I - A^{-1} b' e_n^T) p = x - A^{-1} b$$

We can solve this by the Woodbury formula or by forming an extended system. Either way, we end up with a Newton step solve that is $O(n^2)$ time.

```
[L,U,P] = lu(A);
for k = 1:maxiter
  r = x−U\(L\(P∗b(x(end))));
  c = U\(L\(P∗db(x(end))));
  z = r(end)/(1−c(end));
  p = r−c∗z;
  x = x + p;
end
```