Formulas

- Inner product, norm, angle.
  - Relation between inner product, norm, and angle: $a^T b = \|a\| \, \|b\| \cos \angle(a, b)$.
  - Average value of elements of an n-vector: $\mathbf{avg}(a) = (\mathbf{1}^T a)/n$.
  - Root-mean-square value of an n-vector: $\mathbf{rms}(a) = \|a\|/\sqrt{n}$.
  - Standard deviation of an n-vector: $\mathbf{std}(a) = \|a - \mathbf{avg}(a)\mathbf{1}\|/\sqrt{n}$.
  - Correlation coefficient of two non-constant n-vectors a and b:

$$\rho = \frac{(a - \mathbf{avg}(a)\mathbf{1})^T (b - \mathbf{avg}(b)\mathbf{1})}{\|a - \mathbf{avg}(a)\mathbf{1}\| \, \|b - \mathbf{avg}(b)\mathbf{1}\|}.$$

- Complexity of basic matrix and vector operations (α is a scalar, x and y are n-vectors, A is an $m \times n$ matrix, B is an $n \times p$ matrix).
  - Inner product $x^T y$: $2n - 1$ flops ($\approx 2n$ flops for large n).
  - Vector addition $x + y$: n flops.
  - Scalar-vector multiplication αx: n flops.
  - Scalar-matrix multiplication αA: mn flops.
  - Matrix-vector multiplication Ax: $m(2n - 1)$ flops ($\approx 2mn$ flops for large n).
  - Matrix-matrix multiplication AB: $mp(2n - 1)$ flops ($\approx 2mpn$ flops for large n).
- Pseudo-inverses.
  - Pseudo-inverse of left-invertible matrix A: $A^\dagger = (A^T A)^{-1} A^T$.
  - Pseudo-inverse of right-invertible matrix A: $A^\dagger = A^T (A A^T)^{-1}$.
- Complexity of forward or back substitution with triangular $n \times n$ matrix: n^2 flops.
- Complexity of matrix factorizations.
  - QR factorization of $m \times n$ matrix: $2mn^2$ flops.
  - LU factorization of $n \times n$ matrix: $(2/3)n^3$ flops.

Problem 1. Show that every 2×2 rotation matrix can be written as a product of two reflectors: for every θ, there exist u, v such that

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = (I - 2uu^T)(I - 2vv^T), \qquad \|u\| = \|v\| = 1.$$

Without loss of generality, one can choose $v = (1, 0)$, so it remains to find u. You may find the trigonometric identities $\cos\theta = \cos^2(\theta/2) - \sin^2(\theta/2)$ and $\sin\theta = 2\cos(\theta/2)\sin(\theta/2)$ useful.

Answer for problem 1.

Let $v = (1,0)$, $vv^T = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ $\qquad (u_1\ u_2)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = u_1$

$$(I - 2uu^T)(I - 2vv^T) = I - 2vv^T - 2uu^T + 4uu^Tvv^T$$

$$= I - 2vv^T - 2uu^T + 4(u^Tv)uv^T \qquad \text{Let } u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

$$u^Tv = u_1$$

$$= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} - 2uu^T + 4u_1 uv^T$$

$$= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} - 2\begin{pmatrix} u_1^2 & u_1u_2 \\ u_1u_2 & u_2^2 \end{pmatrix} + 4u_1\begin{pmatrix} u_1 & 0 \\ u_2 & 0 \end{pmatrix} \qquad \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}(u_1\ u_2)$$

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}(1\ 0)$$

$$= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} - 2\begin{pmatrix} u_1^2 & u_1u_2 \\ u_1u_2 & u_2^2 \end{pmatrix} + 4\begin{pmatrix} u_1^2 & 0 \\ u_1u_2 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 2u_1^2 & -2u_1u_2 \\ 2u_1u_2 & -2u_2^2 \end{pmatrix}$$

$$= \begin{pmatrix} -1 + 2u_1^2 & -2u_1u_2 \\ 2u_1u_2 & 1 - 2u_2^2 \end{pmatrix}$$

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2} & -2\cos\frac{\theta}{2}\sin\theta/2 \\ 2\cos\frac{\theta}{2}\sin\frac{\theta}{2} & \cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2} \end{pmatrix}$$

Let $u_1 = \cos\frac{\theta}{2}$, $u_2 = \sin\frac{\theta}{2}$

$$-1 + 2u_1^2 = -1 + 2\cos^2\tfrac{\theta}{2} = -(\cos^2\tfrac{\theta}{2} + \sin^2\tfrac{\theta}{2}) + 2\cos^2\tfrac{\theta}{2}$$
$$= \cos^2\tfrac{\theta}{2} - \sin^2\tfrac{\theta}{2}$$

$$1 - 2u_2^2 = 1 - 2\sin^2\tfrac{\theta}{2} = (\cos^2\tfrac{\theta}{2} + \sin^2\tfrac{\theta}{2}) - 2\sin^2\tfrac{\theta}{2}$$
$$= \cos^2\tfrac{\theta}{2} - \sin^2\tfrac{\theta}{2}$$

With $v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $u = \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix}$, we get

$$(I - 2uu^T)(I - 2vv^T) = \begin{pmatrix} \cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2} & -2\cos\frac{\theta}{2}\sin\frac{\theta}{2} \\ 2\cos\frac{\theta}{2}\sin\frac{\theta}{2} & \cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2} \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Problem 2. Consider the $n \times n$ matrix $A = aI - \mathbf{1}\mathbf{1}^T$ where a is a real scalar and $\mathbf{1}$ is the n-vector of ones. We assume $n > 1$.

1. For what values of a is A singular?
2. Assuming A is nonsingular, express its inverse as a linear combination of I and $\mathbf{1}\mathbf{1}^T$.

$$\mathbf{1}\mathbf{1}^T = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}(1, \ldots, 1) = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix}$$

Answer for problem 2.

$$A = \begin{pmatrix} a & 0 & \cdots & 0 \\ 0 & a & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & a \end{pmatrix} - \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} a-1 & -1 & \cdots & -1 \\ -1 & a-1 & & \vdots \\ \vdots & & \ddots & -1 \\ -1 & \cdots & -1 & a-1 \end{pmatrix}$$

1) A singular when $Ax = 0$ for $x \neq 0$

$$Ax = ax - \mathbf{1}\mathbf{1}^T x = 0$$

$$ax = \mathbf{1}\mathbf{1}^T x \qquad \mathbf{1}\mathbf{1}^T x = (\mathbf{1}^T x)\mathbf{1} = (x_1 + \cdots + x_n)\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} ax_1 \\ \vdots \\ ax_n \end{pmatrix} = \begin{pmatrix} x_1 + \cdots + x_n \\ \vdots \\ x_1 + \cdots + x_n \end{pmatrix}$$

If $a = 0$, $x = (1, 0, \ldots, 0, -1)$ is a solution.

$ax_1 = ax_2 = \cdots = ax_n = x_1 + \cdots + x_n$ only satisfied if $a = 0$ or $x_1 = \cdots = x_n$

If $x_1 = x_2 = \cdots = x_n$, then $a = n$ solves the equation.

A is singular when $a = 0$, or $a = n$

2) $A^{-1} = cI + d\mathbb{1}\mathbb{1}^T$

$$A^{-1}A = (cI + d\mathbb{1}\mathbb{1}^T)(aI - \mathbb{1}\mathbb{1}^T) = I$$

$$= acI + ad\mathbb{1}\mathbb{1}^T - c\mathbb{1}\mathbb{1}^T - d(\mathbb{1}\mathbb{1}^T)(\mathbb{1}\mathbb{1}^T) = I$$

$$= acI + (ad - c)\mathbb{1}\mathbb{1}^T - d\mathbb{1}\mathbb{1}^T\mathbb{1}\mathbb{1}^T$$

$$\mathbb{1}\mathbb{1}^T\mathbb{1}\mathbb{1}^T = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \\ 1 & & 1 \end{pmatrix}\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \\ 1 & \cdots & 1 \end{pmatrix}$$

$$= \begin{pmatrix} n & \cdots & n \\ \vdots & & \\ n & \cdots & n \end{pmatrix}$$

$$= n\mathbb{1}\mathbb{1}^T$$

$$A^{-1}A = acI + (ad - c - nd)\mathbb{1}\mathbb{1}^T = I$$

$$ac = 1 \Rightarrow c = 1/a$$

$$ad - c - nd = 0$$

$$ad - \frac{1}{a} - nd = 0$$

$$(a-n)d = 1/a \Rightarrow d = \frac{1}{a(a-n)}$$

$$A^{-1} = \frac{1}{a}I + \frac{1}{a(a-n)}\mathbb{1}\mathbb{1}^T$$

#

Problem 3. Consider an $m \times m$ block matrix A of the form

$$A = \begin{bmatrix} A_{11} & 0 & 0 & \cdots & 0 & 0 \\ 0 & A_{22} & 0 & \cdots & 0 & 0 \\ 0 & 0 & A_{33} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_{m-1,m-1} & 0 \\ A_{m1} & A_{m2} & A_{m3} & \cdots & A_{m,m-1} & A_{mm} \end{bmatrix}.$$

Each block A_{ij} and all the zero matrices in the definition are matrices of size $n \times n$, so the matrix A itself has size $nm \times nm$. What is the complexity of solving $Ax = b$ in each of the following cases?

1. The matrices A_{ii} are lower triangular and nonsingular.
2. The matrices A_{ii} are orthogonal.
3. The matrices A_{ii} are general nonsingular matrices.

Include in the complexity all terms that are cubic (n^3, n^2m, nm^2, m^3) or higher-order. Explain your answers.

Answer for problem 3.

$Ax = b \quad Ax = \begin{pmatrix} A_{11} & 0 & \cdots & & 0 \\ 0 & A_{22} & & & \vdots \\ \vdots & & \ddots & A_{m-1,m-1} & 0 \\ A_{m1} & \cdots & & & A_{mm} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$ each x_i is an n-vector.

$= \begin{pmatrix} A_{11} x_1 \\ A_{22} x_2 \\ \vdots \\ A_{m-1,m-1} x_{m-1} \\ A_{m1} x_1 + \cdots + A_{mm} x_m \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ Each b_i is an n-vector.

4'

1) $A_{11} x_1 = b_1, \ldots, A_{m-1,m-1} x_{m-1} = b_{m-1}$ can be solved by forward substitution

$\rightarrow (m-1)(n^2) = mn^2 - n^2$ complexity

Then $x_1, \ldots, x_{m-1}$ are given, so $A_{m1} x_1 + \cdots + A_{m,m-1} x_{m-1}$ is matrix-vector mult, and addition

$\rightarrow \underbrace{(m-1)\cdot(2n^2)}_{\text{mult}} + \underbrace{(m-2)n}_{\text{add}} = 2mn^2 - 2n^2 + mn - 2n$ complexity

Then $A_{mm} x_m = b_m - \sum_{i=1}^{m-1} A_{mi} x_i$ can be solved by forward substitution

$\rightarrow n^2$ complexity.

$+n$ vector add

2'

Total complexity: $3mn^2$

2) For $1 \le i \le m-1$, we solve $A_{ii} x_i = b_i$ by computing $x_i = A_{ii}^T b_i$ as A_{ii} is orthogonal.

$\rightarrow (m-1)(2n^2) = 2mn^2 - 2n^2$ complexity

Then $x_1, \dots, x_{m-1}$ are given, so $c = b_m - \sum_{i=1}^{m-1} A_{mi} x_i$ is $m-1$ matrix-vector multiplications and $m-1$ vector adds.

$\rightarrow (m-1)(2n^2) + (m-1)(n) = 2mn^2 - 2n^2 + mn - n$ complexity

Then solve $A_{mm} x_m = c$ by computing $x_m = A_{mm}^T c$

$\rightarrow 2n^2$ complexity.

Total complexity: $4mn^2$ 2'

3) Step 1: Factorize $A_{ii} = P_i L_i U_i$ for $1 \le i \le m$

$\rightarrow m \cdot \frac{2}{3} n^3 = \frac{2}{3} mn^3$

Step 2: For $1 \le i \le m-1$, solve $A_{ii} x_i = b_i$ by $P_i L_i U_i x_i = b_i$

a) Solve $P_i y_i = b_i \rightarrow 0$ flops (permutation only)

b) Solve $L_i z_i = y_i \rightarrow n^2$ flops (forward sub)

c) Solve $U_i x_i = z_i \rightarrow n^2$ flops (back sub)

$\rightarrow (m-1)(2n^2) = 2mn^2 - 2n^2$ flops

Step 3: Compute $c = b_m - \sum_{i=1}^{m-1} A_{mi} x_i$

$\rightarrow (m-1)(2n^2) + (m-1)(n) = 2mn^2 - 2n^2 + mn - n$

Step 4: Solve $A_{mm} x_m = c$ by $P_m L_m U_m x_m = c$

a) Solve $P_m y_m = b_m \rightarrow 0$ flops (permutation)

b) Solve $L_m z_m = y_m \rightarrow n^2$ flops (forward sub)

c) Solve $U_m x_m = z_m \rightarrow n^2$ flops (back sub)

$\rightarrow 2n^2$ flops

Total complexity: $\frac{2}{3} mn^3 + 4mn^2$ 2'

Problem 4. Let A be a nonsingular $n \times n$ matrix. Explain how each of the following problems can be solved with a single LU factorization of A. Describe the steps in your algorithms and give the complexity of each step, including all quadratic and cubic terms in the flop count. If you know several methods, give the most efficient one.

1. Find the $n \times n$ matrix X that satisfies

$$AXA = uv^T$$

where u and v are given n-vectors.

2. Find the n-vectors x, y that satisfy

$$\begin{bmatrix} 0 & A \\ A^T & B \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b \\ c \end{bmatrix}.$$

The $n \times n$ matrix B, and the n-vectors b, c are given.

3. Find the n-vectors x, y that satisfy

$$\begin{bmatrix} A & A \\ -A & A \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b \\ c \end{bmatrix}.$$

The n-vectors b, c are given.

Answer for problem 4.

1) $AXA = uv^T \qquad X = A^{-1}uv^TA^{-1}$ Let $y = A^{-1}u$ $X = yv^TA^{-1}$

$X^T = (A^T)^{-1}vy^T \qquad$ Let $z = (A^T)^{-1}v$

$= zy^T$

$X = yz^T$

Step 1: Factorize $A = PLU \to \frac{2}{3}n^3$ flops

2: Solve $PLUy = u \to 2n^2$ flops

a) Solve $Py_1 = u \to 0$ flops (permutation)

b) Solve $Ly_2 = y_1 \to n^2$ flops (fwd sub)

c) Solve $Uy = y_2 \to n^2$ flops (back sub)

3: Solve $U^TL^TP^Tz = v \to 2n^2$ flops

a) Solve $U^Tz_1 = v \to n^2$ flops (fwd sub)

b) Solve $L^Tz_2 = z_1 \to n^2$ flops (back sub)

c) Solve $P^Tz = z_2 \to 0$ flops (permutation)

4: Compute $X = yz^T \to n^2$ flops

Total complexity: $\frac{2}{3}n^3 + 5n^2$ flops

2) $\begin{bmatrix} 0 & A \\ A^T & B \end{bmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ c \end{pmatrix} \rightarrow \begin{matrix} Ay = b \\ A^T x + By = c \end{matrix}$

Step 1: Factorize $A = PLU$. $\rightarrow \frac{2}{3}n^3$ flops

$A^T = U^T L^T P^T$

2: Solve $PLUy = b \rightarrow 2n^2$ flops

a) Solve $Py_1 = b \rightarrow 0$ flops (permutation)

b) Solve $Ly_2 = y_1 \rightarrow n^2$ flops (forward sub)

c) Solve $Uy = y_2 \rightarrow n^2$ flops (back sub).

3) Compute $By = z \rightarrow 2n^2$ flops

4) Compute $d = c - z \rightarrow n$ flops

5) Solve $U^T L^T P^T x = d \rightarrow 2n^2$ flops

a) Solve $U^T x_1 = d \rightarrow n^2$ flops (forward sub)

b) Solve $L^T x_2 = x_1 \rightarrow n^2$ flops (back sub)

c) Solve $P^T x = x_2 \rightarrow 0$ flops (permutation)

Total complexity: $\frac{2}{3}n^3 + 6n^2$ flops

3) $\begin{pmatrix} A & A \\ -A & A \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ c \end{pmatrix}$ $\quad \begin{matrix} Ax + Ay = b \\ -Ax + Ay = c \end{matrix} \quad \begin{matrix} A(x+y) = b \\ A(-x+y) = c \end{matrix} \quad \begin{matrix} 2Ay = b + c \\ 2Ax = b - c \end{matrix}$

Step 1: Factorize $A = PLU \rightarrow \frac{2}{3}n^3$ flops

2: Compute $d = \frac{1}{2}(b+c) \rightarrow 2n$ flops

3: Compute $f = \frac{1}{2}(b-c) \rightarrow 2n$ flops

4: Solve $PLUy = d \rightarrow 2n^2$ flops

a) Solve $Py_1 = d \rightarrow 0$ flops (permutation)

b) Solve $Ly_2 = y_1 \rightarrow n^2$ flops (forward sub)

c) Solve $Uy = y_2 \rightarrow n^2$ flops (back sub)

5: Solve $PLUx = f \rightarrow 2n^2$ flops

a) Solve $Px_1 = f \rightarrow 0$ flops (permutation)

b) Solve $Lx_2 = x_1 \rightarrow n^2$ flops (forward sub)

c) Solve $Ux = x_2 \rightarrow n^2$ flops (back sub)

Total complexity: $\frac{2}{3}n^3 + 4n$

flops

Problem 5. Suppose A and B are left-invertible matrices of the same size that satisfy

$$AA^T = BB^T.$$

In this problem we show that $B = AQ$ for some orthogonal Q.

1. Show that the matrix $U = A^\dagger B$ is orthogonal.
2. Show that the matrix $V = B^\dagger A$ is orthogonal.
3. Show that U is the inverse of V.
4. Find an orthogonal matrix Q such that $B = AQ$.

Answer for problem 5.

1) $A^+ = (A^TA)^{-1}A^T$

$V = (A^TA)^{-1}A^TB \qquad V^T = B^TA(A^TA)^{-T} = B^TA(A^TA)^{-1}$

$VV^T = ((A^TA)^{-1}A^TB)(B^TA(A^TA)^{-1}) \qquad (A^TA)^{-T} = ((A^TA)^T)^{-1} = (A^TA)^{-1}$

$= (A^TA)^{-1}A^TBB^TA(A^TA)^{-1}$

$= (A^TA)^{-1}A^TAA^TA(A^TA)^{-1}$

$= (A^TA)^{-1}(A^TA)(A^TA)(A^TA)^{-1}$

$VV^T = I$ so V is orthogonal ✓

2) $B^+ = (B^TB)^{-1}B^T$

$V = (B^TB)^{-1}B^TA \qquad V^T = A^TB(B^TB)^{-T} = A^TB(B^TB)^{-1}$

$VV^T = ((B^TB)^{-1}B^TA)(A^TB(B^TB)^{-1})$

$= (B^TB)^{-1}B^TAA^TB(B^TB)^{-1}$

$= (B^TB)^{-1}B^TBB^TB(B^TB)^{-1}$

$VV^T = I$ so V is orthogonal. ✓

3) Show $UV = I$ ✓

$UV = (A^+B)(B^+A) = (A^TA)^{-1}A^TB(B^TB)^{-1}B^TA$

$A^T = A^+BB^T$ by $A^+AA^T = A^+BB^T$

$B^T = B^+AA^T$

$UV = (A^TA)^{-1}A^+BB^TB(B^TB)^{-1}B^TA$

$= (A^TA)^{-1}A^+BB^TA$

$= (A^TA)^{-1}A^+AA^TA$

$= (A^TA)^{-1}A^TA$

$UV = I$

$VU = (B^TB)^{-1}B^TA(A^TA)^{-1}A^TB$

$= (B^TB)^{-1}B^+AA^TA(A^TA)^{-1}A^TB$

$= (B^TB)^{-1}B^+AA^TB$

$= (B^TB)^{-1}B^+BB^TB = (B^TB)^{-1}B^TB$

$VU = I$

4) $B = AQ$

$BB^T = AA^T \qquad B^+BB^T = B^+AA^T$

$B^T = B^+AA^T = (B^TB)^{-1}B^TAA^T$

$B = AA^TB(B^TB)^{-1}$

$= AV^T$ ✓

$Q = V^T$