Special Matrices
(Strict) Diagonal Dominance

- The magnitude of each diagonal element is (either):
  - strictly larger than the sum of the magnitudes of all the other elements in its row
  - strictly larger than the sum of the magnitudes of all the other elements in its column
- One may row/column scale and permute rows/columns to achieve diagonally dominant (since it’s just a rewriting of the equations)
  - Recall: choosing the form of the equations wisely is important

- E.g. consider $\begin{pmatrix} 3 & -2 \\ 5 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 9 \\ 4 \end{pmatrix}$
- Switch rows $\begin{pmatrix} 5 & 1 \\ 3 & -2 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 9 \end{pmatrix}$ and column scale $\begin{pmatrix} 5 & -2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} -.5c_2 \\ c_1 \end{pmatrix} = \begin{pmatrix} 4 \\ 9 \end{pmatrix}$
(Strict) Diagonal Dominance

- **Strictly** diagonally dominant (square) matrices are guaranteed to be non-singular.
- Since $\det(A) = \det(A^T)$, either row or column diagonal dominance is enough.
- Column diagonal dominance guarantees that pivoting is not required during $LU$ factorization.
- However, pivoting still improves robustness.
- E.g. consider $\begin{pmatrix} 4 & 3 \\ -2 & 50 \end{pmatrix}$ where 50 is more desirable than 4 for $a_{11}$. 
Inner Product

- Consider the space of all vectors with length $m$
- The dot/inner product of two vectors is $u \cdot v = \sum_i u_i v_i$
- The magnitude of a vector is $\|v\|_2 = \sqrt{v \cdot v} \geq 0$
- Alternative notations: $\langle u, v \rangle = u \cdot v = u^T v$

- Weighted inner product defined via an $n \times n$ matrix $A$
- $\langle u, v \rangle_A = u \cdot Av = u^T Av$
- Since $\langle v, u \rangle_A = v^T Au = u^T A^T v$, weighted inner products commute when $A$ is symmetric
- The standard dot product uses identity matrix weighting: $\langle u, v \rangle = \langle u, v \rangle_I$
Definiteness

• Assume A is symmetric so that $< u, v >_A = < v, u >_A$

• $A$ is **positive definite** if and only if $< v, v >_A = v^T A v > 0$ for $\forall v \neq 0$
• $A$ is **positive semi-definite** if and only if $< v, v >_A = v^T A v \geq 0$ for $\forall v \neq 0$
• We abbreviate with SPD and SP(S)D

• $A$ is **negative definite** if and only if $< v, v >_A = v^T A v < 0$ for $\forall v \neq 0$
• $A$ is **negative semi-definite** if and only if $< v, v >_A = v^T A v \leq 0$ for $\forall v \neq 0$
• If $A$ is negative (semi) definite, then $-A$ is positive (semi) definite (and vice versa)
• Thus, can convert such problems to SPD or SP(S)D

• $A$ is considered **indefinite** when it is neither positive/negative semi-definite
Eigenvalues

• SPD matrices have all eigenvalues \( > 0 \)
• SP(S)D matrices have all eigenvalues \( \geq 0 \)

• Symmetric negative definite matrices have all eigenvalues \( < 0 \)
• Symmetric negative semi-definite matrices have all eigenvalues \( \leq 0 \)

• Indefinite matrices have both positive and negative eigenvalues
Recall: SVD Construction (Unit 3)

• Let \( A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \) so that \( A^T A = AA^T = I \), and thus \( U = V = \Sigma = I \)

• But \( A \neq U\Sigma V^T = I \) What’s wrong?

• Given a column vector \( v_k \) of \( V \), \( Av_k = U\Sigma V^T v_k = U\Sigma \hat{e}_k = U\sigma_k \hat{e}_k = \sigma_k u_k \) where \( u_k \) is the corresponding column of \( U \)

\[
Av_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = u_1 \text{ but } Av_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 1 \end{pmatrix} = u_2
\]

• Since \( U \) and \( V \) are orthonormal, their columns are unit length

• However, there are still two choices for the direction of each column

• Multiplying \( u_2 \) by \(-1\) to get \( u_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \) makes \( U = A \), and thus \( A = U\Sigma V^T \) as desired
Symmetric Matrices (SVD)

• Since $A^T A = AA^T = A^2$, both the columns of $U$ and the columns of $V$ are eigenvectors of $A^2$

• They are identical (but potentially opposite) directions: $u_k = \pm v_k$

• Thus, $A v_k = \sigma_k u_k$ implies $A v_k = \pm \sigma_k v_k$

• That is, the $v_k$ (and $u_k$) are eigenvectors of $A$ with eigenvalues $\pm \sigma_k$

• Similar to the polar SVD, can pull negative signs out of the columns of $U$ into the $\sigma_k$ to obtain $U = V$ and $A = V \Lambda V^T$ as a modified SVD

• $A = V \Lambda V^T$ implies $AV = V \Lambda$ which is the matrix form of the eigensystem of $A$

• Here, $\Lambda$ contains the positive and negative eigenvalues of $A$
SPD Matrices

• When $A$ is SP(S)D, $\Lambda = \Sigma$ and the standard SVD is $A = V\Sigma V^T$ (i.e. $U = V$)

• The singular values are the (all positive) eigenvalues of $A$ (since $AV = V\Sigma$)

• Constructing $V$ with $\det V = 1$ (as usual) and obtaining all $\sigma_k > 0$ implies that there are no reflections

• Since all $\sigma_k > 0$, SPD matrices have full rank and are invertible

• SP(S)D (and not SPD) has at least one $\sigma_k = 0$ and a null space

• Often, one can use modified SPD techniques for SP(S)D matrices

• Unfortunately, indefinite matrices are significantly more challenging
Making/Breaking Symmetry

• Row/column scaling can break symmetry:
  • Row scaling \( \begin{pmatrix} 5 \\ 3 \\ -4 \end{pmatrix} \) by \(-2\) gives a non-symmetric \( \begin{pmatrix} 5 \\ -6 \\ 3 \\ 8 \end{pmatrix} \)
  • Additional column scaling by \(-2\) gives \( \begin{pmatrix} 5 \\ -6 \\ -6 \\ -16 \end{pmatrix} \)
  • Scaling the same row/column together in the same way preserves symmetry

• Important: a nonsymmetric matrix might be inherently symmetric when properly rescaled/rearranged
Rules Galore

• There are many rules/theorems regarding special matrices (especially for SPD)
• It is important to be aware of reference material (and to look things up)

• Examples:
  • SPD matrices don’t require pivoting during $LU$ factorization
  • A symmetric (strictly) diagonally dominant matrix with positive diagonal entries is positive definite
  • Jacobi and Gauss-Seidel iteration converge when a matrix is strictly (or irreducibly) diagonally dominant
  • Etc.
Cholesky Factorization

• SPD matrices have $LU$ factorization of $LL^T$ and don’t require elimination to find it

• Consider

$$\begin{pmatrix} a_{11} & a_{21} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{pmatrix} \begin{pmatrix} l_{11} & l_{21} \\ 0 & l_{22} \end{pmatrix} = \begin{pmatrix} l_{11}^2 & l_{11}l_{21} \\ l_{21}l_{11} & l_{21}^2 + l_{22}^2 \end{pmatrix}$$

• So $l_{11} = \sqrt{a_{11}}$ and $l_{21} = \frac{a_{21}}{l_{11}}$ and $l_{22} = \sqrt{a_{22} - l_{21}^2}$

```plaintext
for(j=1,n){
    for(k=1,j-1) for(i=j,n) a_{ij} -= a_{ik}a_{jk};
    a_{jj} = \sqrt{a_{jj}}; for(k=j+1,n) a_{kj} /= a_{jj};
}
```

\ For each column j of the matrix
\ For each column j of the matrix
\ Loop over all previous columns k, and subtract a multiple of column k from the current column j
\ Take the square root of the diagonal entry, and scale column j by that value

• This algorithm factors the matrix “in place” replacing $A$ with $L$
Incomplete Cholesky Preconditioner

• Cholesky factorization can be used to construct a preconditioner for a sparse matrix.
• The full Cholesky factorization would fill in too many non-zero entries.
• So, *incomplete* Cholesky preconditioning uses Cholesky factorization with the *caveat* that only the nonzero entries are modified (all zeros remain zeros).
Symmetric Approximation

- For non-symmetric $A$, a symmetric $\hat{A} = \frac{1}{2} (A + A^T)$ averages off-diagonal components
- Solving the symmetric $\hat{A}c = b$ instead of the non-symmetric $Ac = b$ gives a faster/easier (perhaps erroneous) approximation to a problem that might not require too much accuracy
- Alternatively, the inverse of the symmetric $\hat{A}$ (or the notion thereof) may be used to devise a preconditioner for $Ac = b$