LU FACTORIZATION OF STRUCTURED MATRICES

The special structure of a matrix can be exploited to obtain more efficient factorizations. Evaluation of the linear combination $\boldsymbol{A}\boldsymbol{x} = x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n$ requires mn floating point operations (flops) for $\boldsymbol{A} \in \mathbb{C}^{m\times n}$. Evaluation of p linear combinations $\boldsymbol{A}\boldsymbol{X}$, $\boldsymbol{X} \in \mathbb{C}^{n\times p}$ requires mnp flops. If it is possible to evaluate $\boldsymbol{A}\boldsymbol{x}$ with fewer operations, the matrix is said to be structured. Examples include:

- Banded matrices $\boldsymbol{A} = [a_{ij}]$, $a_{ij} = 0$ if $i - j > l$ or $j - i > u$, with l, u denoting the lower and upper bandwidths. If $l = u = 0$ the matrix is diagonal. If $l = u = b$ the matrix is said to have bandwidth $B = 2b + 1$, i.e., for $b = 1$, the matrix is tridiagonal, and for $b = 2$ the matrix is pentadiagonal. Lower triangular matrices have $u = 0$, while upper triangular matrices have $l = 0$. The $\boldsymbol{A}\boldsymbol{x}$ product requires $(l + u + 1)m$ flops.
- Sparse matrices have r non-zero elements per row or c non-zero elements per column. The $\boldsymbol{A}\boldsymbol{x}$ product requires rm or cn flops
- Circulant matrices $\boldsymbol{A} = [a_{ij}]$ are sqaure and have $a_{ij} = f(i - j)$, a property that can be exploited to compute $\boldsymbol{A}\boldsymbol{x}$ using $\mathcal{O}(m \log m)$ operations
- For square, rank-deficient matrices $\boldsymbol{A} \in \mathbb{C}^{m\times m}$, rank $(\boldsymbol{A}) = r$, $\boldsymbol{A}\boldsymbol{x}$ can be evaluated in $\mathcal{O}(km)$ flops
- When $\boldsymbol{A}, \boldsymbol{X}$ are symmetric (hence square), $\boldsymbol{A}\boldsymbol{X}$ requires $\mathcal{O}(m^3/2)$ flops instead of m^3.

1. Cholesky factorization of positive definite hermitian matrices

1.1. Symmetric matrices, hermitian matrices

Special structure of a matrix is typically associated with underlying symmetries of a particular phenomenon. For example, the law of action and reaction in dynamics (Newton's third law) leads to real symmetric matrices, $\boldsymbol{A} \in \mathbb{R}^{m\times m}$, $\boldsymbol{A}^T = \boldsymbol{A}$. Consider a system of m point masses with nearest-neighbor interactions on the real line where the interaction force depends on relative position. Assume that the force exerted by particle $i + 1$ on particle i is linear

$$f_{i+1,i} = f(u_{i+1} - u_i) = k(u_{i+1} - u_i),$$

with u_i denoting displacement from an equilibrium position. The law of action and reaction then states that

$$f_{i,i+1} = -f_{i+1,i} = k(u_i - u_{i+1}).$$

If the same force law holds at all positions, then

$$f_{i-1,i} = k(u_{i-1} - u_i).$$

The force on particle i is given by the sum of forces from neighboring particles $i - 1, i + 1$

$$f_i = f_{i-1,i} + f_{i+1,i} = k(u_{i-1} - u_i) + k(u_{i+1} - u_i) = k(u_{i+1} - 2u_i + u_{i-1}).$$

Introducing $\boldsymbol{f}, \boldsymbol{u} \in \mathbb{R}^m$, and assuming $u_0 = u_{m+1} = 0$, the above is stated as

$$\boldsymbol{f} = \boldsymbol{K}\boldsymbol{u},$$

with $\boldsymbol{K} = k\,\mathrm{diag}([\ 1\ \ -2\ \ 1\ ])$ is a symmetric matrix, $\boldsymbol{K} = \boldsymbol{K}^T$, a direct consequence of the law of action and reaction. The matrix $\boldsymbol{K}$ is in this case tridiagonal as a consequence of the assumption of nearest-neighbor interactions. Recall that matrices represent linear mappings, hence

$$\boldsymbol{K} = [\ \boldsymbol{f}(\boldsymbol{e}_1)\ \ \boldsymbol{f}(\boldsymbol{e}_2)\ \ \ldots\ \ \boldsymbol{f}(\boldsymbol{e}_m)\ ],$$

with $\boldsymbol{f}(\boldsymbol{u})$ the force-displacement linear mapping, Fig. 1, obtaining the same symmetric, tri-diagonal matrix.

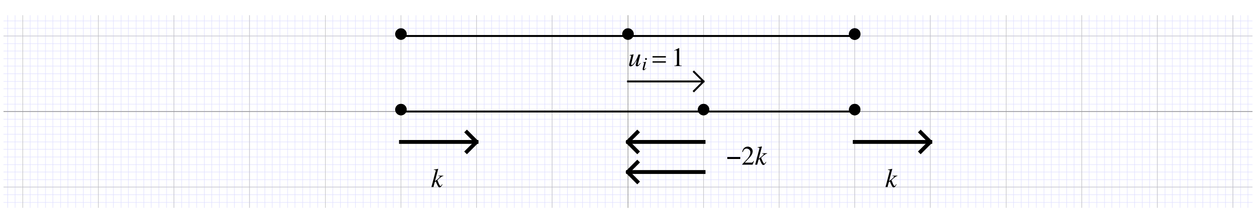

Figure 1. Image of $\boldsymbol{e}_i$ through mapping representing a linear force is $\boldsymbol{f}(\boldsymbol{e}_i) = k[\ \ldots\ \ 1\ \ -2\ \ 1\ \ \ldots\ ]^T$.

This concept can be extended to complex matrices $\boldsymbol{A} \in \mathbb{C}^{m \times m}$ through $\boldsymbol{A}^* = \boldsymbol{A}$, in which case $\boldsymbol{A}$ is said to be self-adjoint or hermitian. Again, this property is often associated with desired physical properties, such as the requirement of real observable quantitites in quantum mechanics. Diagonal elements of a hermitian matrix must be real, and for any $\boldsymbol{x}$, $\boldsymbol{y} \in \mathbb{C}^m$, the computation

$$\overline{\boldsymbol{x}^* \boldsymbol{A} \boldsymbol{y}} = (\boldsymbol{x}^* \boldsymbol{A} \boldsymbol{y})^* = \boldsymbol{y}^* \boldsymbol{A}^* \boldsymbol{x} = \boldsymbol{y}^* \boldsymbol{A} \boldsymbol{x},$$

implies for $\boldsymbol{x} = \boldsymbol{y}$ that

$$\overline{\boldsymbol{x}^* \boldsymbol{A} \boldsymbol{x}} = \boldsymbol{x}^* \boldsymbol{A} \boldsymbol{x},$$

hence $\boldsymbol{x}^* \boldsymbol{A} \boldsymbol{x}$ is real.

1.2. Positive-definite matrices

The work (i.e., energy stored in the system) done by all the forces in the above point mass system is

$$\mathscr{W} = \frac{1}{2} \boldsymbol{u}^T \boldsymbol{K} \boldsymbol{u},$$

and physical considerations state that $\mathscr{W} \geqslant 0$. This leads the following definitions.

DEFINITION. *A hermitian matrix* $\boldsymbol{A} \in \mathbb{C}^{m \times m}$ *is positive definite if for any non-zero* $\boldsymbol{x} \in \mathbb{C}^m$, $\boldsymbol{x}^* \boldsymbol{A} \boldsymbol{x} > 0$.

DEFINITION. *A hermitian matrix* $\boldsymbol{A} \in \mathbb{C}^{m \times m}$ *is positive semi-definite if for any non-zero* $\boldsymbol{x} \in \mathbb{C}^m$, $\boldsymbol{x}^* \boldsymbol{A} \boldsymbol{x} \geqslant 0$.

If $\boldsymbol{A}$ is hermitian positive definite, then so is $\boldsymbol{X}^* \boldsymbol{A} \boldsymbol{X}$ for any $\boldsymbol{X} \in \mathbb{C}^{m \times n}$. Choosing

$$\boldsymbol{X} = [\ \boldsymbol{e}_1 \ \ldots \ \boldsymbol{e}_n\ ] \in \mathbb{C}^{m \times n}$$

gives $\boldsymbol{A}_n = \boldsymbol{X}^* \boldsymbol{A} \boldsymbol{X}$, the n^{th} principal submatrix of $\boldsymbol{A}$, itself a hermitian positive definite matrix. Choosing $\boldsymbol{X} = \boldsymbol{e}_j$shows that the j^{th} diagonal element of $\boldsymbol{A}$ is positive, $a_{jj} = \boldsymbol{e}_j^T \boldsymbol{A} \boldsymbol{e}_j > 0$

1.3. Symmetric factorization of positive-definite hermitian matrices

The structure of a hermitian positive definite matrix $\boldsymbol{A} \in \mathbb{C}^{m \times m}$, can be preserved by modification of LU-factorization. The resulting algorithm is known as Cholesky factorization, and its first stage is stated as

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & \boldsymbol{w}^* \\ \boldsymbol{w} & \boldsymbol{B} \end{bmatrix} = \begin{bmatrix} \alpha & \boldsymbol{0} \\ \boldsymbol{w}/\alpha & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} 1 & \boldsymbol{0}^* \\ \boldsymbol{0} & \boldsymbol{C} \end{bmatrix} \begin{bmatrix} \alpha & \boldsymbol{w}^*/\alpha \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} = \begin{bmatrix} \alpha & \boldsymbol{0} \\ \boldsymbol{w}/\alpha & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \alpha & \boldsymbol{w}^*/\alpha \\ \boldsymbol{0} & \boldsymbol{C} \end{bmatrix} = \begin{bmatrix} a_{11} & \boldsymbol{w}^* \\ \boldsymbol{w} & \boldsymbol{C} + \boldsymbol{w}\boldsymbol{w}^*/a_{11} \end{bmatrix},$$

whence $\boldsymbol{C} = \boldsymbol{B} - \boldsymbol{w}\boldsymbol{w}^*/a_{11}$. Repeating the stage-1 step

$$\boldsymbol{A} = \boldsymbol{L}_1 \boldsymbol{A}_1 \boldsymbol{L}_1^*,$$

leads to

$$\boldsymbol{A} = \boldsymbol{L}_1 \boldsymbol{L}_2 \boldsymbol{A}_2 \boldsymbol{L}_2^* \boldsymbol{L}_1^* = \cdots = \boldsymbol{L}\boldsymbol{L}^*, \boldsymbol{L} = \boldsymbol{L}_1 \boldsymbol{L}_2 \ldots \boldsymbol{L}_m.$$

The resulting Cholesky algorithm is half as expensive as standard LU-factorization.

Algorithm (Cholesky factorization, $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}^*$)

```
L = A
for i = 1:m
  for j = i+1:m
    L[j:m, j] = L[j:m, j] − L[j:m, i] L̄[j, i] / L[i, i]
  L[i:m, i] = L[i:m, i] / √L[i, i]
```

2. iLU-factorization of sparse matrices

The two-dimensional extension of the nearest-neighbor interacting point mass system