Review and catch up

- Frequently Used Matrix Decompositions
- What else in the Handout

Frequently Used Matrix Decompositions

"Big four":

1. LU factorization

   Gaussian elimination, ...

2. QR decomposition

   Gram-Schmidt orthogonalization, ...

3. Schur decomposition

   Eigendecomposition, ...

4. Singular value decomposition

   CS decomposition, ...

LU decomposition (Gaussian Elimination)

If A is nonsingular, then there exist permutations P, a unit lower triangular matrix L, and a nonsingular upper triangular matrix U such that

$$PA = LU.$$

Special cases:

(a) **Cholesky decomposition.** A matrix A is symmetric positive definite <u>if and only if</u> there exists a unique nonsingular upper triangular matrix R, with positive diagonal entries, such that

$$A = R^T R.$$

(b) **LDL$^{\mathrm{T}}$ factorization** If $A^T = A$ is nonsingular, then there exists a permutation P, a unit lower triangular matrix L, and a block diagonal matrix D with 1-by-1 and 2-by-2 blocks such that

$$PAP^T = LDL^T.$$

Applications:

- Solve $Ax = b$.
- ...

Matlab demo: `lugui.m`, `bslashtx.m`

QR decomposition(Gram-Schmidt orthogonalization)

Let A be m-by-n with $m \geq n$. Suppose that A has full column rank. Then there exist a unique m-by-n orthogonal matrix Q ($Q^T Q = I$) and a unique n-by-n upper triangular matrix R with positive diagonal $r_{ii} > 0$ such that

$$A = QR.$$

Applications:

- Find an orthonormal basis of the subspace spanned by the columns of A.
- Solve the linear least squares problem $\min_x \|Ax - b\|_2$.

Matlab demo: `qrsteps.m`

Schur decomposition

Let A be of order n. Then there is an $n \times n$ unitary matrix U ($U^H U = I$) such that

$$A = UTU^H,$$

where T is upper triangular. By appropriate choice of U, the eigenvalues of A, which are the diagonal elements of T, may be made to appear in any order.

Applications:

- Compute eigenvalues and eigenvectors of A.
- ...

Matlab demo: `eigsvdgui.m`

Singular Value Decomposition (SVD)

Let A be an m-by-n matrix with $m \geq n$. Then we can write

$$A = U\Sigma V^T,$$

where U is m-by-n orthogonal matrix ($U^TU = I_n$) and V is n-by-n orthogonal matrix ($V^TV = I$), and $\Sigma = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$, where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$.

If $m < n$, the SVD can be defined by considering A^T.

The columns $u_1, u_2, \ldots, u_n$ of U are called *left singular vectors* of A. The columns $v_1, v_2, \ldots, v_n$ of V are called *right singular vectors*. The $\sigma_1, \sigma_2, \ldots, \sigma_n$ are called singular values.

Applications:

- Suppose that A is m-by-n with $m \geq n$ and has full rank, with $A = U\Sigma V^T$ being A's SVD. Then the pseudo-inverse can also be written as

$$A^\dagger \equiv (A^T A)^{-1} A^T = V \Sigma^{-1} U^T.$$

  (If $m < n$, then $A^\dagger = A^T(AA^T)^{-1}$)

- Suppose that

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0,$$

  Then the rank of A is r. The range space of A is span$(u_1, u_2, \cdots, u_r)$. and the null space of A is span$(v_{r+1}, v_{r+2}, \ldots, v_n)$.

- $\|A\|_2 = \sigma_1 (\equiv \sigma_{\max})$

- Let A be $m \times n$ with $m \geq n$. Then

  (a) eigenvalues of $A^T A = V\Sigma^2 V^T$ are σ_i^2, $i = 1, 2, \ldots, n$. The corresponding eigenvectors are the right singular vectors v_i, $i = 1, 2, \ldots, n$.

  (b) eigenvalues of $AA^T = U\Sigma^2 U^T$ are σ_i^2, $i = 1, 2, \ldots, n$ and $m - n$ zeros. The left singular vectors u_i, $i = 1, 2, \ldots, n$ are corresponding eigenvectors for the eigenvalues σ_i^2. One can take any $m - n$ other orthogonal vectors that are orthogonal to $u_1, u_2, \ldots, u_n$ as the eigenvectors for the eigenvalues 0.

- Optimal rank-k approximation:

$$\min_{\substack{B : m \times n \\ \text{rank}(B) = k}} \|A - B\|_2 = \|A - A_k\|_2 = \sigma_{k+1},$$

where $A_k = U\Sigma_k V^T$, $\Sigma_k =$ diag$(\sigma_1, \sigma_2, \ldots, \sigma_k, 0, \ldots, 0)$

Note that A_k can be written in a compact form as

$$A_k = U_k \widehat{\Sigma}_k V_k^T,$$

where U_k and V_k are the first k columns of U and V, respectively, $\widehat{\Sigma}_k =$ diag$(\sigma_1, \sigma_2, \ldots, \sigma_k)$. Therefore, A_k is represented by $mk + k + nk = (m + n + 1)k$ elements, in contrast, A is represented by mn elements.

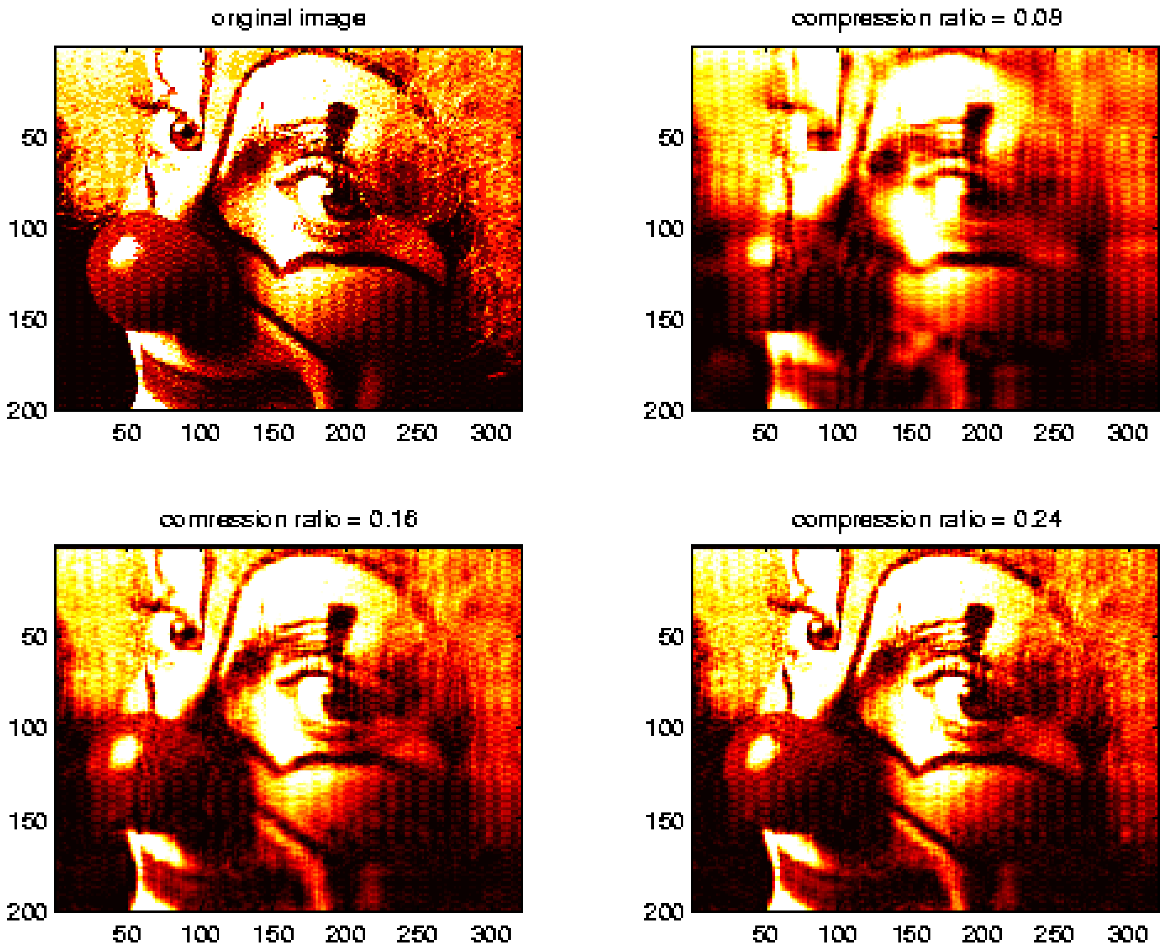
Application: image compression.

$$\text{compression ratio} = \frac{(m + n + 1)k}{mn}$$

Matlab’s M-scripts:

```
>> load clown.mat;
>> [m,n]=size(X);
>> figure(1);
>> colormap(map);
>> imag(X);
>> [U,S,V]=svd(X);
>> k = 20;
>> X20 = U(:,1:k)*S(1:k,1:k)*V(:,1:k)';
>> figure(2);
>> colormap(map);
>> image(X20);
>> compression_ratio = (m+n)*k/(m*n)
```

Output:


original image
compression ratio = 0.08
comression ratio = 0.16
compression ratio = 0.24
50
100
150
200
50
100
150
200
250
300

What else in the Handout

1. Approximation and error are the facts of life in matrix computing
2. Floating-point arithmetic
3. Rounding error analysis
4. Vector and matrix norms
5. The landscape of linear systems solvers
6. A general framework for iterative projection methods for solving $Ax = b$
   - Steepest descent $\Rightarrow$ Conjugate Gradient (CG) method
   - Minimal residual $\Rightarrow$ GMRES
7. The concept of preconditioning
8. A general framework for iterative projection methods for solving $Ax = \lambda x$
   - Arnoldi method
   - Lanczos methods