SYMMETRIC POSITIVE DEFINITE (SPD) MATRICES SPD LINEAR SYSTEMS

- Symmetric positive definite matrices.
- The LDL^T decomposition; The Cholesky factorization

Positive-Definite Matrices

A real matrix is said to be positive definite if

$$(Au,u)>0$$
 for all $u
eq 0$ $u\in \ \mathbb{R}^n$

Let A be a real positive definite matrix. Then there is a scalar $\alpha > 0$ such that

$$(Au,u)\geq lpha\|u\|_2^2.$$

Consider now the case of Symmetric Positive Definite (SPD) matrices.

$$\blacktriangleright$$
 Consequence 1: A is nonsingular

6-2

• Consequence 2: the eigenvalues of A are (real) positive

GvL 4 – SPD

A few properties of SPD matrices

 \blacktriangleright Diagonal entries of A are positive

> Recall: the k-th principal submatrix A_k is the $k \times k$ submatrix of A with entries $a_{ij}, 1 \leq i, j \leq k$ (Matlab: A(1:k, 1:k)).

 \blacksquare_1 Each A_k is SPD

6-3

Consequence: $Det(A_k) > 0$ for $k = 1, \dots, n$. In fact A is SPD iff this condition holds.

11 If A is SPD then for any $n \times k$ matrix X of rank k, the matrix $X^T A X$ is SPD.

► The mapping : $x, y \rightarrow (x, y)_A \equiv (Ax, y)$ defines a proper inner product on \mathbb{R}^n . The associated norm, denoted by $\|\cdot\|_A$, is called the energy norm, or simply the A-norm:

$$\|x\|_A = (Ax, x)^{1/2} = \sqrt{x^T A x}$$

Related measure in Machine Learning, Vision, Statistics: the Mahalanobis distance between two vectors:

$$d_A(x,y) = \|x-y\|_A = \sqrt{(x-y)^T A(x-y)}$$

Appropriate distance (measured in # standard deviations) if x is a sample generated by a Gaussian distribution with covariance matrix A and center y.

More terminology

A matrix is Positive Semi-Definite if:

6-5

 $(Au,u)\geq 0$ for all $u\in \mathbb{R}^n$

Eigenvalues of symmetric positive semi-definite matrices are real nonnegative, i.e., ...

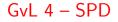
> ... A can be singular [If not, A is SPD]

> A matrix is said to be Negative Definite if -A is positive definite. Similar definition for Negative Semi-Definite

► A matrix that is neither positive semi-definite nor negative semidefinite is indefinite

And $(Ax, x) = 0 \; orall x$ then A = 0

Show: $A \neq 0$ is indefinite iff $\exists x, y : (Ax, x)(Ay, y) < 0$



The LDL^{T} and Cholesky factorizations

\swarrow_{6} The (standard) LU factorization of an SPD matrix A exists

 \blacktriangleright Let A = LU and D = diag(U) and set $M \equiv (D^{-1}U)^T$.

Then
$$A = LU = LD(D^{-1}U) = LDM^T$$

Both L and M are unit lower triangular

► Consider $L^{-1}AL^{-T} = DM^TL^{-T}$

6-6

> Matrix on the right is upper triangular. But it is also symmetric. Therefore $M^T L^{-T} = I$ and so M = L > Alternative proof: exploit uniqueness of LU factorization without pivoting + symmetry: $A = LDM^T = MDL^T \rightarrow M = L$

The diagonal entries of D are positive [Proof: consider $L^{-1}AL^{-T} = D$]. In the end:

$$A = LDL^T = GG^T$$
 where $G = LD^{1/2}$

Cholesky factorization is a specialization of the LU factorization for the SPD case. Several variants exist.

6-7

First algorithm: row-oriented LDLT

Adapted from Gaussian Elimination. Main observation: The working matrix A(k+1:n,k+1:n) in standard LU remains symmetric. \rightarrow Work only on its upper triangular part & ignore lower part

1. For
$$k = 1 : n - 1$$
 Do:
2. For $i = k + 1 : n$ Do:
3. $piv := a(k,i)/a(k,k)$
4. $a(i,i:n) := a(i,i:n) - piv * a(k,i:n)$
5. End
6. End

This will give the U matrix of the LU factorization. Therefore D = diag(U), $L^T = D^{-1}U$.



Row-Cholesky (outer product form)

6-9

Scale the rows as the algorithm proceeds. Line 4 becomes

$$a(i,:):=a(i,:)\!-\![a(k,i)/\sqrt{a(k,k)}]*\left[a(k,:)/\sqrt{a(k,k)}
ight]$$

ALGORITHM : 1. Outer product Cholesky

1. For
$$k = 1 : n$$
 Do:
2. $A(k, k : n) = A(k, k : n) / \sqrt{A(k, k)}$;
3. For $i := k + 1 : n$ Do :
4. $A(i, i : n) = A(i, i : n) - A(k, i) * A(k, i : n)$;
5. End
6. End

> Result: Upper triangular matrix U such $A = U^T U$.

Example:

$$A=egin{pmatrix} 1 & -1 & 2\ -1 & 5 & 0\ 2 & 0 & 9 \end{pmatrix}$$

 \swarrow_7 Is A symmetric positive definite?

- Multiple What is the LDL^T factorization of A?
- Mhat is the Cholesky factorization of A ?

GvL 4 – SPD

Column Cholesky. Let $A = GG^T$ with G = lower triangular. Then equate j-th columns:

$$a(:,j) = \sum_{k=1}^{j} g(:,k) g^{T}(k,j)
ightarrow$$

$$egin{aligned} A(:,j) &= \sum\limits_{k=1}^{j} G(j,k) G(:,k) \ &= G(j,j) G(:,j) + \sum\limits_{k=1}^{j-1} G(j,k) G(:,k)
ightarrow \ G(j,j) G(:,j) &= A(:,j) - \sum\limits_{k=1}^{j-1} G(j,k) G(:,k) \end{aligned}$$

GvL 4 – SPD

6-11

> Assume that first j-1 columns of G already known.

Compute unscaled column-vector:

$$v = A(:,j) - \sum_{k=1}^{j-1} G(j,k) G(:,k)$$

> Notice that $v(j) \equiv G(j,j)^2$.

6-12

 \blacktriangleright Compute $\sqrt{v(j)}$ and scale v to get j-th column of G.

ALGORITHM : 2 Column Cholesky

1. For
$$j = 1 : n \ do$$

2. For $k = 1 : j - 1 \ do$
3. $A(j:n,j) = A(j:n,j) - A(j,k) * A(j:n,k)$
4. EndDo
5. If $A(j,j) \le 0$ ExitError("Matrix not SPD")
6. $A(j,j) = \sqrt{A(j,j)}$
7. $A(j+1:n,j) = A(j+1:n,j)/A(j,j)$
8. EndDo

Try algorithm on:

$$A=egin{pmatrix} 1 & -1 & 2 \ -1 & 5 & 0 \ 2 & 0 & 9 \end{pmatrix}$$

6-13

GvL 4 – SPD