Fall-23 5304 LecN7 Notes

Wan

November 9, 2023

Topics: The Gram-Schmidt algo and the QR Factorization; Least-squares problems; Applications; Data fitting.

Math tools: Eigenvalues and Eigencevectors

1 Least-Squares Systems and Data fitting

Problem BG

1, overdetermined system, more euqations than unknowns, no sol

如果一个方程组无解，那么这个方程组被称为不一致。例如下面的方程组：

$x_1 + x_2 = 2$
$x_1 - x_2 = 1$
$x_1 + x_2 = 3$

根据线性代数$^\mathbb{Q}$的知识，m个方程n个未知量 $m > n$ 时通常无解，但是虽然不能求出 $Ax = b$ 的解，那何不退而求其次，去寻找与解近似的向量 x 。

那么如何定义与解相似，一般使用欧氏距离来进行度量，即两点间的距离，这其实很好理解，越相似，欧氏距离越近，这样求出的 x 被称为最小二乘解。

2, Function approximation
   3, Goal: Find the best approximation to the system of equations.

Geometric Interpretation

Good illustration: Data fitting

2 Gram-Schmidt

2.1 Concept

2.2 Classical Gram-Schmidt

The cost of classical Gram-Schmidt is $2mn^2 - \frac{2}{3}n^3$ flops.

Verification of Algo:
1, A = QR. by checking whether norm(A - Q*R) is a small number.
2, Orthogonality: Q is orthogonal. by checking whether $Q^T Q = I$, norm(Q'* Q - eye(n)) should be small.

2.3 Modified Gram-Schmidt

Cost of QR by using classical Gram-Schmidt:

Cost

3 QR Factorization

3.1 BG

The norm here is referring to the 2-norm.

3.2 Concept

> **Definition 8.6 QR-factorization**
>
> Let A be an $m \times n$ matrix with independent columns. A **QR-factorization** of A expresses it as $A = QR$ where Q is $m \times n$ with orthonormal columns and R is an invertible and upper triangular matrix with positive diagonal entries.

The importance of the factorization lies in the fact that there are computer algorithms that accomplish it with good control over round-off error, making it particularly useful in matrix calculations. The factorization is a matrix version of the Gram-Schmidt process.

A = QR where Q is an orthogonal matrix and R is an invertible upper triangular matrix. If A is non singular, then this factorization is unique.

3.3 Process

Suppose $A = [c_1, c_2, ...c_n]$ is a full-rank matrix, size $m \times n$.

Basically 3 steps:
1, find A's orthogonal set $F = [f_1, f_2...f_n]$.
2, normalize each col of F to get $Q = [q_1, q_2...q_n]$.
3, find R by $R = Q^T A$.

To find F, use this formula where $f_1 = c_1$ (based on the orthogonal set):
$f_k = c_k - \frac{c_k * f_1}{\|f_1\|^2} f_1 - \frac{c_k * f_2}{\|f_2\|^2} f_2 - ... - \frac{c_k * f_{k-1}}{\|f_{k-1}\|^2} f_{k-1}$ (k > 2)

To find Q, normalize each col of F by $q_j = \frac{f_j}{\|f_j\|}$, j = 1 to n.
At this time, $q_1, q_2...q_n$ are orthonormal cols.

QR Visulization (c_j can be expressed as a linear combination of q_j):

Using these equations, express each $\mathbf{c}_k$ as a linear combination of the $\mathbf{q}_i$:

$$
\begin{aligned}
\mathbf{c}_1 &= \|\mathbf{f}_1\|\mathbf{q}_1 \\
\mathbf{c}_2 &= (\mathbf{c}_2\cdot\mathbf{q}_1)\mathbf{q}_1 + \|\mathbf{f}_2\|\mathbf{q}_2 \\
\mathbf{c}_3 &= (\mathbf{c}_3\cdot\mathbf{q}_1)\mathbf{q}_1 + (\mathbf{c}_3\cdot\mathbf{q}_2)\mathbf{q}_2 + \|\mathbf{f}_3\|\mathbf{q}_3 \\
&\vdots \qquad \vdots \\
\mathbf{c}_n &= (\mathbf{c}_n\cdot\mathbf{q}_1)\mathbf{q}_1 + (\mathbf{c}_n\cdot\mathbf{q}_2)\mathbf{q}_2 + (\mathbf{c}_n\cdot\mathbf{q}_3)\mathbf{q}_3 + \cdots + \|\mathbf{f}_n\|\mathbf{q}_n
\end{aligned}
$$

These equations have a matrix form that gives the required factorization:

$$
\begin{aligned}
A &= \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 & \cdots & \mathbf{c}_n \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 & \cdots & \mathbf{q}_n \end{bmatrix}
\begin{bmatrix}
\|\mathbf{f}_1\| & \mathbf{c}_2\cdot\mathbf{q}_1 & \mathbf{c}_3\cdot\mathbf{q}_1 & \cdots & \mathbf{c}_n\cdot\mathbf{q}_1 \\
0 & \|\mathbf{f}_2\| & \mathbf{c}_3\cdot\mathbf{q}_2 & \cdots & \mathbf{c}_n\cdot\mathbf{q}_2 \\
0 & 0 & \|\mathbf{f}_3\| & \cdots & \mathbf{c}_n\cdot\mathbf{q}_3 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & \|\mathbf{f}_n\|
\end{bmatrix}
\end{aligned}
\tag{8.5}
$$

Here the first factor $Q = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 & \cdots & \mathbf{q}_n \end{bmatrix}$ has orthonormal columns, and the second factor is an $n \times n$ upper triangular matrix R with positive diagonal entries (and so is invertible). We record this in the following theorem.

Examples

See topic Notes.

Cost

lalala

Application

QR分解主要有三个用途：

- （1）首先它可以用于求解n个方程n个未知变量的系统 $Ax = b$，如果A非奇异，那么就可以将A分解为QR，回代则求解出x，但是QR分解线性方程组🔍比LU分解方法的计算代价大三倍还多，用的不多。
- （2）用于求解最小二乘问题🔍
- （3）用于特征值计算，这个之后的帖子再写

Application1: Use QR to solve least-squares problems

11-09

　　unitary matrix perserves the length of a vector.

　　rank-1 update
　　3 steps:
1, transform problem 2, gradient