ĐẠI HỌC BÁCH KHOA HÀ NỘI
TRƯỜNG CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

***** *****

BÁO CÁO BÀI TẬP LỚN

HỌC PHẦN: TÍNH TOÁN KHOA HỌC – IT4110

Chủ đề:

PHÂN RÃ QR VÀ BÀI TOÁN BÌNH PHƯƠNG TỐI THIỂU

Giảng viên hướng dẫn: **TS. Tạ Duy Hoàng**

Nhóm sinh viên thực hiện: **Nhóm 10**

STT	Họ và tên	MSSV
1	Ngô Trọng Dũng	202416896
2	Trần Văn Cường	202416872
3	Đỗ Tùng Long	202416972
4	Nguyễn Trọng Linh	202416968
5	Lê Đặng Thành An	
6	Phan Gia Phong	
7	Vũ Trần Tuấn Minh	

MỤC LỤC

4. CHƯƠNG 4: CÀI ĐẶT VÀ THỰC NGHIỆM

4.1 Môi trường và công cụ sử dụng

4.2 Xây dựng chương trình giải LS bằng QR

4.3 Kiểm thử với dữ liệu thực tế

4.4 Phân tích sai số

4.5 Đánh giá hiệu quả thuật toán

5. CHƯƠNG 5: KẾT LUẬN

6. TÀI LIỆU THAM KHẢO

MỞ ĐẦU

Trong thực tế, dữ liệu thu thập từ các thí nghiệm hoặc quá trình quan sát thường chịu ảnh hưởng của sai số đo và nhiễu, khiến cho các mô hình toán học khó có thể mô tả chính xác hoàn toàn mối quan hệ giữa các đại lượng. Do đó, việc xây dựng một mô hình xấp xỉ phản ánh tốt nhất xu hướng của dữ liệu đóng vai trò quan trọng và xuất hiện phổ biến trong nhiều lĩnh vực như toán học ứng dụng, kỹ thuật và khoa học dữ liệu.

Bài toán bình phương tối thiểu (Least Squares Problem) được đề xuất nhằm giải quyết vấn đề xấp xỉ dữ liệu trong trường hợp hệ phương trình tuyến tính quá định, thông qua việc tối thiểu hóa tổng bình phương sai số giữa mô hình và dữ liệu thực nghiệm. Với tính chất cơ bản và hiệu quả, phương pháp này đã trở thành một công cụ quan trọng trong đại số tuyến tính và phân tích số liệu.

Trong số các phương pháp giải bài toán bình phương tối thiểu, phương pháp phân rã QR được sử dụng rộng rãi nhờ ưu điểm về độ ổn định số, hạn chế ảnh hưởng của sai số làm tròn và khả năng tính toán hiệu quả đối với các hệ phương trình kích thước lớn. Vì vậy, việc nghiên cứu và áp dụng phân rã QR để giải bài toán bình phương tối thiểu có ý nghĩa cả về mặt lý thuyết lẫn thực tiễn.

Xuất phát từ những lý do trên, báo cáo này được thực hiện nhằm tìm hiểu bài toán bình phương tối thiểu tuyến tính và phương pháp phân rã QR, đồng thời áp dụng phương pháp này để giải một số bài toán cụ thể thông qua cài đặt và thực nghiệm minh họa. Nội dung báo cáo gồm năm chương: Chương 1 trình bày tổng quan về bài toán bình phương tối thiểu; Chương 2 giới thiệu phương pháp phân rã QR; Chương 3 trình bày cách giải bài toán bình phương tối thiểu bằng phân rã QR; Chương 4 tập trung vào cài đặt và thực nghiệm; và Chương 5 đưa ra kết luận và hướng phát triển tiếp theo.

CHƯƠNG 1: BÀI TOÁN BÌNH PHƯƠNG TỐI THIỂU

1.1. Giới thiệu bài toán

1.1.1. Bối cảnh và động cơ của bài toán

Trong nhiều bài toán thực tế thuộc các lĩnh vực khoa học, kỹ thuật và phân tích dữ liệu, dữ liệu thu thập được thường xuất phát từ các phép đo hoặc quan sát thực nghiệm. Do chịu ảnh hưởng của sai số đo lường, nhiễu hoặc các yếu tố ngẫu nhiên khác, dữ liệu này hiếm khi thỏa mãn chính xác một mô hình toán học lý tưởng. Vì vậy, thay vì tìm một nghiệm chính xác, mục tiêu đặt ra là tìm một mô hình xấp xỉ sao cho sai lệch giữa mô hình và dữ liệu là nhỏ nhất theo một tiêu chí nhất định. Bài toán bình phương tối thiểu (Least Squares Problem) được xây dựng nhằm đáp ứng yêu cầu đó và đóng vai trò quan trọng trong đại số tuyến tính ứng dụng và phương pháp số.

1.1.2. Ví dụ minh họa: bài toán hồi quy tuyến tính

Một ví dụ điển hình của bài toán bình phương tối thiểu là bài toán hồi quy tuyến tính. Giả sử ta có một tập dữ liệu gồm n điểm thực nghiệm $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, trong đó các giá trị x_iđược giả định là chính xác, còn các giá trị y_i bị ảnh hưởng bởi nhiễu. Ta mong muốn tìm một đường thẳng có dạng

$$y = \beta_0 + \beta_1 x$$

sao cho đường thẳng này mô tả dữ liệu một cách tốt nhất. Do ảnh hưởng của nhiễu, đường thẳng này không thể đi qua tất cả các điểm dữ liệu, nhưng vẫn được xem là mô hình xấp xỉ hợp lý cho tập dữ liệu đã cho.

1.1.3. Hệ phương trình quá định

Khi viết các phương trình tương ứng với từng điểm dữ liệu trong bài toán hồi quy tuyến tính, ta thu được một hệ gồm n phương trình với hai ẩn β_0 và β_1. Đây là một hệ phương trình quá định, trong đó số phương trình lớn hơn số ẩn. Một cách tổng quát, hệ này có thể được viết dưới dạng phương trình ma trận

$$Ax = b,$$

với $A \in \mathbb{R}^{m \times n}$ và $m > n$. Trong trường hợp này, do vector $b$thường không thuộc không gian cột của ma trận A, hệ phương trình không có nghiệm chính xác. Hiện tượng này là đặc trưng phổ biến của các bài toán xuất phát từ dữ liệu thực nghiệm, chứ không phải là sai sót trong việc xây dựng mô hình.

1.1.4. Ý tưởng cơ bản của bài toán bình phương tối thiểu

Thay vì tìm nghiệm chính xác của phương trình $Ax = b$, bài toán bình phương tối thiểu tìm một vector $x$sao cho Ax gần với $b$nhất theo chuẩn Euclid. Cụ thể, bài toán được phát biểu dưới dạng tìm $x$sao cho giá trị của biểu thức

$$\|Ax - b\|_2$$

đạt giá trị nhỏ nhất. Nghiệm thu được theo cách này được gọi là nghiệm bình phương tối thiểu. Theo quan điểm hình học, nghiệm bình phương tối thiểu tương ứng với hình chiếu trực giao của vector b lên không gian cột của ma trận A.

1.1.5. Vai trò và ý nghĩa của bài toán

Bài toán bình phương tối thiểu có vai trò nền tảng trong nhiều lĩnh vực như hồi quy thống kê, xử lý tín hiệu, học máy và các bài toán xấp xỉ số. Ngoài ý nghĩa lý thuyết, bài toán này còn là cơ sở cho việc phát triển các phương pháp tính toán hiệu quả và ổn định trong thực hành, tiêu biểu là các phương pháp dựa trên phân rã ma trận như phân rã QR. Do đó, việc nghiên cứu bài toán bình phương tối thiểu không chỉ giúp làm rõ bản chất toán học của vấn đề mà còn tạo tiền đề cho các phương pháp giải quyết hiệu quả trong các chương tiếp theo của báo cáo.

1.2. Cơ sở lý thuyết

Mục này trình bày các khái niệm toán học nền tảng cần thiết để nghiên cứu bài toán bình phương tối thiểu. Các nội dung được lựa chọn nhằm làm rõ bản chất đại số và hình học của bài toán, đồng thời tạo cơ sở cho việc xây dựng mô hình toán học ở Mục 1.3 và phương pháp giải bằng phân rã QR ở các chương tiếp theo.

1.2.1. Không gian vector và không gian con

Một không gian vector trên trường số thực $\mathbb{R}$là một tập hợp các vector cùng với hai phép toán là cộng vector và nhân với vô hướng, thỏa mãn các tiên đề quen thuộc của đại số tuyến tính. Trong báo cáo này, ta chủ yếu làm việc với không gian $\mathbb{R}^n$.

Một tập con $W \subset \mathbb{R}^n$được gọi là không gian con nếu nó đóng đối với phép cộng và phép nhân vô hướng. Mọi tổ hợp tuyến tính của các vector trong Wđều thuộc về W.

Khái niệm không gian con đóng vai trò quan trọng trong bài toán bình phương tối thiểu, bởi nghiệm của bài toán được xác định thông qua mối quan hệ giữa vector dữ liệu và một không gian con xác định bởi ma trận hệ số [1].

1.2.2. Không gian cột của ma trận

Cho ma trận $A \in \mathbb{R}^{m\times n}$, không gian cột của A, ký hiệu là $Col(A)$, được định nghĩa là tập hợp tất cả các tổ hợp tuyến tính của các cột của A:

$$Col(A) = \{Ax \mid x \in \mathbb{R}^n\}.$$

Không gian cột là một không gian con của $\mathbb{R}^m$. Phương trình tuyến tính

$$Ax = b$$

có nghiệm khi và chỉ khi $b \in Col(A)$. Trong trường hợp hệ phương trình là quá định $(m > n)$, điều kiện này thường không được thỏa mãn, do đó hệ không có nghiệm chính xác.

Bài toán bình phương tối thiểu xuất phát trực tiếp từ thực tế này: thay vì tìm nghiệm chính xác của $Ax = b$, ta tìm một vector trong $Col(A)$"gần" với $b$nhất theo một tiêu chí đo lường phù hợp [1], [2].

1.2.3. Tích vô hướng và chuẩn Euclid

Trong không gian $\mathbb{R}^n$, tích vô hướng của hai vector $u, v \in \mathbb{R}^n$được định nghĩa bởi

$$\langle u, v \rangle = u^T v.$$

Từ tích vô hướng, ta xác định chuẩn Euclid của một vector $v$là

$$\| v \|_2 = \sqrt{\langle v, v \rangle}.$$

Chuẩn Euclid cung cấp một cách đo "độ lớn" của vector và được sử dụng rộng rãi để đo sai số trong các bài toán xấp xỉ. Trong bài toán bình phương tối thiểu, tiêu chí tối ưu là làm nhỏ nhất bình phương chuẩn Euclid của sai số $\|Ax - b\|_2^2$, tương đương với việc tối thiểu hóa tổng bình phương sai số trên từng thành phần .

1.2.4. Trực giao và hình chiếu trực giao

Hai vector $u, v \in \mathbb{R}^n$được gọi là trực giao nếu

$$\langle u, v \rangle = 0.$$

Khái niệm trực giao được mở rộng cho các không gian con. Cho Wlà một không gian con của $\mathbb{R}^n$, mỗi vector $b \in \mathbb{R}^n$có thể được phân tích thành tổng của hai thành phần:

$$b = b_W + b_{W^\perp},$$

trong đó $b_W \in W$ và $b_{W^\perp}$trực giao với mọi vector trong W.

Vector b_W được gọi là hình chiếu trực giao của b lên không gian con W. Trong bài toán bình phương tối thiểu, nghiệm thu được chính là vector $x$sao cho Axlà hình chiếu trực giao của $b$lên không gian cột của A. Diễn giải hình học này giúp làm rõ bản chất của bài toán và là nền tảng cho các phương pháp giải ổn định về mặt số học [1], [3].

1.2.5. Một số nhận xét về sự tồn tại nghiệm trong bài toán xấp xỉ

Trong đại số tuyến tính, nhiều bài toán thực tế không yêu cầu tìm nghiệm chính xác của một hệ phương trình tuyến tính, mà quan tâm đến việc tìm một nghiệm **xấp xỉ tối ưu** theo một tiêu chí cho trước. Sự tồn tại của nghiệm trong các bài toán xấp xỉ này gắn liền với cấu trúc của không gian vector và các không gian con liên quan.

Đối với một không gian con $W \subset \mathbb{R}^n$, mọi vector trong $\mathbb{R}^n$đều có thể được xấp xỉ bằng một vector thuộc W theo nghĩa hình chiếu trực giao. Điều này đảm bảo rằng, trong các bài toán xấp xỉ dựa trên chuẩn Euclid, luôn tồn tại ít nhất một nghiệm tối ưu. Tuy nhiên, việc nghiệm đó có duy nhất hay không còn phụ thuộc vào các tính chất đại số của không gian con được xét.

Trong bối cảnh các bài toán tuyến tính, nếu một không gian con được sinh bởi một tập vector độc lập tuyến tính, thì phép chiếu trực giao là duy nhất. Ngược lại, khi các vector sinh phụ thuộc tuyến tính, có thể tồn tại nhiều nghiệm xấp xỉ khác nhau cùng đạt giá trị tối ưu. Những nhận xét này đóng vai trò nền tảng cho việc nghiên cứu các bài toán xấp xỉ và sẽ được áp dụng cụ thể cho bài toán bình phương tối thiểu trong mục tiếp theo.

1.3. Phát biểu và mô hình toán học của bài toán bình phương tối thiểu

Mục này trình bày cách phát biểu bài toán bình phương tối thiểu dưới dạng toán học, đồng thời làm rõ bản chất hình học và đại số của bài toán. Từ đó, ta xây dựng mô hình toán học làm cơ sở cho việc nghiên cứu các phương pháp giải trong các chương tiếp theo.

1.3.1. Phát biểu bài toán bình phương tối thiểu

Cho ma trận $A \in \mathbb{R}^{m\times n}$với $m > n$và vector $b \in \mathbb{R}^m$. Trong nhiều trường hợp thực tế, hệ phương trình tuyến tính

$$Ax = b$$

là một hệ quá định và không có nghiệm do vector $b$không thuộc không gian cột của A.

Bài toán bình phương tối thiểu được phát biểu như sau:
tìm vector $x \in \mathbb{R}^n$sao cho chuẩn Euclid của sai số

$$r = Ax - b$$

đạt giá trị nhỏ nhất, tức là

$$\min_{x\in\mathbb{R}^n} \| Ax - b \|_2.$$

Nghiệm của bài toán trên được gọi là nghiệm bình phương tối thiểu. Cách phát biểu này xuất hiện một cách tự nhiên trong các bài toán xấp xỉ dữ liệu và hồi quy tuyến tính, và là dạng chuẩn được trình bày trong các giáo trình đại số tuyến tính hiện đại.

1.3.2. Diễn giải hình học của bài toán

Dưới góc nhìn hình học, bài toán bình phương tối thiểu có thể được hiểu thông qua hình chiếu trực giao trong không gian Euclid. Cụ thể, tập tất cả các vector có dạng Axchính là không gian cột của ma trận A, ký hiệu $Col(A)$, là một không gian con của $\mathbb{R}^m$.

Do $b \notin Col(A)$, ta tìm một vector $p \in Col(A)$sao cho khoảng cách giữa $b$và $p$là nhỏ nhất. Khi đó, $p$chính là hình chiếu trực giao của $b$lên $Col(A)$. Nếu tồn tại x^{*}sao cho

$$Ax^{*} = p,$$

thì x^{*}là nghiệm bình phương tối thiểu của bài toán.

Điều kiện hình học tương đương là vector sai số

$$r = b - Ax^{*}$$

phải trực giao với không gian cột của A. Diễn giải này đóng vai trò trung tâm trong lý thuyết least squares và được trình bày rõ ràng trong các tài liệu của Strang và Trefethen.

1.3.3. Phương trình chuẩn (Normal Equations)

Từ điều kiện trực giao của sai số với không gian cột, ta có:

$$A^T(b - Ax) = 0.$$

Biến đổi lại, ta thu được hệ phương trình

$$A^T Ax = A^T b,$$

được gọi là phương trình chuẩn của bài toán bình phương tối thiểu.

Hệ phương trình này là một hệ $n$phương trình với nẩn và luôn có nghiệm nếu $A^T A$khả nghịch. Phương trình chuẩn cung cấp một cách tiếp cận đại số để tìm nghiệm bình phương tối thiểu và thường được sử dụng trong các trình bày lý thuyết cơ bản [1], [2].

Tuy nhiên, trong thực hành tính toán số, việc giải phương trình chuẩn có thể gặp vấn đề về sai số làm tròn, đặc biệt khi ma trận Acó điều kiện số kém.

1.3.4. Điều kiện tồn tại và tính duy nhất của nghiệm

Bài toán bình phương tối thiểu luôn tồn tại ít nhất một nghiệm, do hàm mục tiêu $\| Ax - b \|_2^2$là hàm lồi và liên tục trên $\mathbb{R}^n$. Tuy nhiên, tính duy nhất của nghiệm phụ thuộc vào hạng của ma trận A.

Nếu các cột của Ađộc lập tuyến tính, hay tương đương với việc Acó hạng đầy đủ cột, thì ma trận $A^T A$là khả nghịch và nghiệm bình phương tối thiểu là duy nhất. Ngược lại, nếu các cột của A phụ thuộc tuyến tính, bài toán có vô số nghiệm, và người ta thường quan tâm đến nghiệm có chuẩn Euclid nhỏ nhất [1], [3].

Những điều kiện này có ý nghĩa quan trọng trong việc lựa chọn phương pháp giải và đánh giá tính ổn định của nghiệm.

1.3.5. Nhận xét về phương pháp giải số

Mặc dù phương trình chuẩn cung cấp một mô hình toán học rõ ràng cho bài toán bình phương tối thiểu, nhưng trong thực hành tính toán số, phương pháp này không phải lúc nào cũng tối ưu do có thể làm tăng điều kiện số của bài toán. Vì lý do đó, các phương pháp giải dựa trên phân rã trực giao, đặc biệt là phân rã QR, thường được ưu tiên sử dụng.

Phân rã QR cho phép giải bài toán bình phương tối thiểu một cách ổn định hơn, đồng thời tránh được việc hình thành trực tiếp ma trận $A^T A$. Đây chính là động cơ dẫn đến việc nghiên cứu phân rã QR trong chương tiếp theo và ứng dụng của nó trong việc giải bài toán bình phương tối thiểu ở Chương 3.

CHƯƠNG 2: PHÂN RÃ QR

Trong nhiều bài toán số học và ứng dụng thực tế, đặc biệt là các bài toán liên quan đến hệ phương trình tuyến tính và xấp xỉ dữ liệu, việc xử lý trực tiếp một ma trận tổng quát thường gặp nhiều khó khăn về mặt tính toán và độ ổn định số. Do đó, các kỹ thuật phân rã ma trận đóng vai trò quan trọng trong việc đơn giản hóa bài toán và cải thiện hiệu quả tính toán. Một trong những kỹ thuật phổ biến và quan trọng nhất là phân rã QR.

2.1 Khái niệm phân rã QR

Phân rã QR của một ma trận $A \in \mathbb{R}^{m\times n}$(với $m \geq n$) là việc biểu diễn ma trận đó dưới dạng tích của hai ma trận

$$A = QR,$$

trong đó $Q \in \mathbb{R}^{m\times n}$là ma trận có các cột trực chuẩn (thỏa mãn $Q^T Q = I$) và $R \in \mathbb{R}^{n\times n}$là ma trận tam giác trên. Trong trường hợp phân rã QR đầy đủ, Qcó thể là ma trận trực giao kích thước $m \times m$, tuy nhiên trong thực hành tính toán, dạng QR "mỏng" (thin QR) thường được sử dụng do tiết kiệm chi phí tính toán và bộ nhớ.

Ý nghĩa quan trọng của phân rã QR nằm ở việc chuyển một bài toán tuyến tính tổng quát về một bài toán đơn giản hơn. Cụ thể, khi thay thế Abằng tích QR, nhiều bài toán liên quan đến Acó thể được quy về việc xử lý ma trận tam giác R, vốn dễ giải hơn rất nhiều so với ma trận ban đầu. Đặc biệt, trong bài toán bình phương tối thiểu, phân rã QR cho phép tránh việc sử dụng hệ phương trình chuẩn $A^T Ax = A^T b$, vốn có thể gây mất ổn định số do khuếch đại sai số làm tròn.

Về mặt hình học, phân rã QR phản ánh quá trình xây dựng một cơ sở trực chuẩn cho không gian cột của ma trận A. Các cột của ma trận Qtạo thành một cơ sở trực chuẩn của không gian cột này, trong khi ma trận Rchứa thông tin về cách các vector ban đầu của Ađược biểu diễn theo cơ sở đó. Nhờ đặc tính trực giao của Q, các phép biến đổi liên quan đến Qkhông làm thay đổi chuẩn Euclid của vector, từ đó giúp cải thiện độ ổn định trong các phép tính số.

Chính vì những lý do trên, phân rã QR được xem là công cụ nền tảng trong đại số tuyến tính số và được sử dụng rộng rãi trong các lĩnh vực như hồi quy tuyến tính, xử lý tín hiệu, học máy và khoa học dữ liệu. Trong các mục tiếp theo, báo cáo sẽ trình bày các phương pháp xây dựng phân rã QR và phân tích ưu, nhược điểm của từng phương pháp trong thực hành tính toán.

2.2. Các phương pháp phân rã QR

Như đã trình bày ở Mục 2.1, phân rã QR cho phép biểu diễn một ma trận Adưới dạng tích $A = QR$, trong đó Qlà ma trận trực giao và Rlà ma trận tam giác trên. Vấn đề đặt ra trong thực hành là làm thế nào để tính toán hiệu quả và ổn định hai ma trận này. Trong mục này, báo cáo trình bày bốn phương pháp phổ biến để xây dựng phân rã QR: Gram–Schmidt cổ điển, Gram–Schmidt cải tiến, biến đổi Householder và phép quay Givens.

2.2.1. Phương pháp Gram–Schmidt cổ điển

Phương pháp Gram–Schmidt cổ điển (Classical Gram–Schmidt – CGS) là cách tiếp cận trực tiếp và dễ hiểu nhất để xây dựng phân rã QR. Ý tưởng cơ bản của phương pháp này là trực giao hóa các cột của ma trận Anhằm tạo ra một hệ vector trực chuẩn.

Giả sử ma trận $A = [a_1, a_2, \dots, a_n]$với các cột $a_i \in \mathbb{R}^m$. Thuật toán Gram–Schmidt cổ điển được thực hiện theo các bước:

1. Chuẩn hóa cột đầu tiên:

$$q_1 = \frac{a_1}{\| a_1 \|}, r_{11} = \| a_1 \|.$$

2. Với mỗi cột $a_j (j \geq 2)$, loại bỏ các thành phần song song với các vector $q_1, \dots, q_{j-1}$:

$$v_j = a_j - \sum_{i=1}^{j-1} (q_i^T a_j) q_i.$$

3. Chuẩn hóa v_jđể thu được q_j, đồng thời các hệ số chiếu $q_i^T a_j$tạo thành các phần tử của ma trận R.

Ví dụ, nếu A là ma trận gồm hai cột gần song song, phương pháp Gram–Schmidt vẫn tạo được Qvà Rtheo đúng lý thuyết. Tuy nhiên, trong tính toán số, các phép trừ giữa những vector gần nhau có thể gây mất mát chữ số có nghĩa, dẫn đến các vector $q_i$không còn trực giao chính xác. Điều này khiến CGS trở nên kém ổn định khi xử lý dữ liệu thực nghiệm hoặc ma trận có điều kiện kém .

Nhận xét:

Gram–Schmidt cổ điển phù hợp cho mục đích minh họa lý thuyết và giảng dạy, nhưng hiếm khi được sử dụng trực tiếp trong các thư viện tính toán số chuyên nghiệp.

2.2.2. Phương pháp Gram–Schmidt cải tiến
Để cải thiện tính ổn định số của Gram–Schmidt cổ điển, người ta sử dụng Gram–Schmidt cải tiến (Modified Gram–Schmidt – MGS). Về mặt toán học, MGS cho kết quả phân rã tương đương CGS, nhưng thứ tự các phép toán được sắp xếp lại nhằm giảm ảnh hưởng của sai số làm tròn.

Trong MGS, thay vì chiếu a_j lên toàn bộ các vector q_i đã có trong một bước, ta lần lượt:

- lấy $a_j^{(0)} = a_j$,
- sau đó cập nhật tuần tự

$$a_j^{(i)} = a_j^{(i-1)} - (q_i^T a_j^{(i-1)}) q_i,$$

- cuối cùng chuẩn hóa $a_j^{(j-1)}$ để thu được q_j.

Cách cập nhật dần này giúp giảm hiện tượng hủy số trong các phép trừ và làm cho các vector thu được gần trực giao hơn trong thực tế tính toán. Theo Lambert, MGS thường cho kết quả ổn định hơn CGS khi áp dụng vào bài toán bình phương tối thiểu, đặc biệt với dữ liệu có nhiễu hoặc các cột của A gần phụ thuộc tuyến tính .

Nhận xét:
MGS là một cải tiến quan trọng và thường được xem như "phiên bản thực dụng" của Gram–Schmidt, nhưng vẫn chưa đạt mức ổn định cao nhất so với Householder.

2.2.3. Biến đổi Householder
Phương pháp Householder sử dụng các phép phản xạ trực giao để đưa ma trận A về dạng tam giác trên. Thay vì trực giao hóa từng cột, phương pháp này triệt tiêu toàn bộ các phần tử dưới đường chéo của mỗi cột trong một bước.

Ý tưởng chính là xây dựng các ma trận phản xạ H_k sao cho:

$$H_k A^{(k)} = \begin{pmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & * \end{pmatrix},$$

với các phần tử dưới đường chéo của cột thứ k bị triệt tiêu. Sau một dãy các phản xạ Householder, ta thu được:

$$R = H_n \cdots H_1 A, Q^T = H_n \cdots H_1.$$

Theo Lambert, phương pháp Householder có tính ổn định số rất cao, vì mỗi phép phản xạ là một biến đổi trực giao hoàn chỉnh, không làm khuếch đại sai số. Do đó, Householder là phương pháp được ưu tiên trong các thư viện tính toán chuẩn như LAPACK khi giải bài toán QR và bình phương tối thiểu .

Nhận xét:
Householder là lựa chọn "chuẩn công nghiệp" cho phân rã QR trong các bài toán kích thước lớn và yêu cầu độ chính xác cao.

2.2.4. Phép quay Givens
Phép quay Givens là một phương pháp khác để xây dựng phân rã QR bằng cách sử dụng các phép quay trong mặt phẳng hai chiều nhằm triệt tiêu từng phần tử riêng lẻ của ma trận.

Mỗi phép quay Givens được thiết kế để làm cho một phần tử cụ thể dưới đường chéo trở thành 0, trong khi chỉ ảnh hưởng đến hai hàng của ma trận. Bằng cách áp dụng một chuỗi các phép quay Givens, ma trận Adần được biến đổi thành ma trận tam giác trên R.

Ưu điểm lớn của phương pháp này là tính linh hoạt: nó đặc biệt hiệu quả đối với ma trận thưa hoặc các bài toán mà ma trận được cập nhật từng phần (ví dụ thêm hoặc xóa một hàng). Lambert chỉ ra rằng trong những trường hợp này, Givens có thể tiết kiệm đáng kể chi phí tính toán so với Householder .

Nhận xét:
Givens phù hợp cho các bài toán cần xử lý cục bộ hoặc online, nhưng kém hiệu quả hơn Householder với ma trận dày và kích thước lớn.

2.3. So sánh các phương pháp phân rã QR
Các phương pháp phân rã QR được trình bày trong Mục 2.2 đều cho phép biểu diễn một ma trận A dưới dạng $A = QR$. Tuy nhiên, chúng khác nhau đáng kể về độ chính xác, tính ổn định số và chi phí tính toán. Trong mục này, các phương pháp Gram–Schmidt cổ điển, Gram–Schmidt cải tiến, Householder và Givens được so sánh dựa trên các tiêu chí quan trọng trong tính toán số, đặc biệt trong bối cảnh giải bài toán bình phương tối thiểu.

2.3.1. Độ chính xác
Độ chính xác của phân rã QR thường được đánh giá thông qua mức độ mà ma trận Qthu được thỏa mãn tính trực giao, tức là mức độ gần với điều kiện $Q^T Q = I$. Trong các phép tính số, sai số làm tròn có thể khiến các vector trong Qkhông còn trực giao hoàn toàn, từ đó ảnh hưởng đến nghiệm của bài toán bình phương tối thiểu.

Phương pháp Gram–Schmidt cổ điển thường cho kết quả kém chính xác nhất trong số các phương pháp được xét, đặc biệt khi các cột của ma trận Agần phụ thuộc tuyến tính. Trong những trường hợp này, hiện tượng hủy số xảy ra mạnh, làm cho các vector trực chuẩn bị sai lệch đáng kể.

Gram–Schmidt cải tiến cho độ chính xác tốt hơn so với phiên bản cổ điển nhờ thay đổi thứ tự các phép toán. Tuy nhiên, khi ma trận có điều kiện kém, phương pháp này vẫn có thể tạo ra sai lệch đáng kể trong tính trực giao của Q.

Ngược lại, các phương pháp dựa trên biến đổi trực giao toàn phần như Householder và Givens thường đạt độ chính xác cao hơn. Đặc biệt, Householder cho phép duy trì tính trực giao của Qgần với sai số máy, do đó thường được xem là phương pháp chính xác nhất trong thực hành tính toán số.

2.3.2. Tính ổn định số

Tính ổn định số là tiêu chí then chốt khi đánh giá các thuật toán phân rã QR, đặc biệt trong các bài toán có dữ liệu nhiễu hoặc ma trận có điều kiện kém. Một thuật toán được xem là ổn định nếu sai số làm tròn không bị khuếch đại đáng kể trong quá trình tính toán.

Theo các tài liệu kinh điển về đại số tuyến tính số, Gram–Schmidt cổ điển là phương pháp kém ổn định nhất do sử dụng nhiều phép trừ giữa các vector gần song song. Gram–Schmidt cải tiến cải thiện đáng kể tính ổn định, nhưng vẫn chưa đạt mức ổn định cao nhất.

Các phương pháp Householder và Givens đều sử dụng các phép biến đổi trực giao, vốn không làm thay đổi chuẩn Euclid của vector. Nhờ đó, hai phương pháp này có tính ổn định số cao. Trong đó, Householder thường được ưu tiên hơn trong các bài toán quy mô lớn vì khả năng xử lý toàn cục và giảm tích lũy sai số. Givens, mặc dù ổn định, nhưng do áp dụng từng phép quay cục bộ nên thường được sử dụng trong các bài toán đặc thù như ma trận thưa hoặc cập nhật từng phần.

2.3.3. Độ phức tạp thuật toán

Về mặt chi phí tính toán, các phương pháp phân rã QR có độ phức tạp khác nhau. Đối với một ma trận $A \in \mathbb{R}^{m\times n}$với $m \geq n$:

- Gram–Schmidt cổ điển và Gram–Schmidt cải tiến có độ phức tạp xấp xỉ $O(mn^2)$, với chi phí tính toán tương đối thấp và thuật toán đơn giản.
- Householder cũng có độ phức tạp $O(mn^2)$, nhưng với hằng số lớn hơn do các phép biến đổi trên toàn bộ cột.

- Givens thường có chi phí cao hơn trong trường hợp ma trận dày, do cần nhiều phép quay để triệt tiêu các phần tử dưới đường chéo. Tuy nhiên, đối với ma trận thưa, phương pháp này có thể tận dụng cấu trúc thưa để giảm đáng kể chi phí tính toán.

Như vậy, xét thuần túy về độ phức tạp, các phương pháp có cùng bậc, nhưng hiệu quả thực tế phụ thuộc mạnh vào cấu trúc của ma trận và bối cảnh ứng dụng.

2.3.4. Nhận xét và khuyến nghị sử dụng

Từ các phân tích trên, có thể rút ra một số nhận xét tổng quát như sau:

- Gram–Schmidt cổ điển phù hợp cho mục đích minh họa lý thuyết, nhưng không nên sử dụng trong các bài toán số yêu cầu độ chính xác cao.
- Gram–Schmidt cải tiến là một lựa chọn tốt hơn khi cần thuật toán đơn giản và dễ cài đặt, với độ ổn định chấp nhận được trong nhiều trường hợp.
- Householder là phương pháp được khuyến nghị cho các bài toán bình phương tối thiểu quy mô lớn nhờ độ chính xác và tính ổn định số cao.
- Givens thích hợp cho các bài toán đặc thù như xử lý ma trận thưa, cập nhật dữ liệu theo thời gian hoặc khi cần triệt tiêu từng phần tử riêng lẻ.

Trong các thư viện tính toán số hiện đại, phân rã QR bằng Householder thường được sử dụng như lựa chọn mặc định, trong khi Givens được dùng trong các tình huống yêu cầu tính linh hoạt cao.

CHƯƠNG 3: GIẢI BÀI TOÁN BÌNH PHƯƠNG TỐI THIỂU BẰNG PHÂN RÃ QR

3.1. Phát biểu lại bài toán bình phương tối thiểu và ý tưởng sử dụng phân rã QR

Xét bài toán bình phương tối thiểu dưới dạng

$$\min_{x \in \mathbb{R}^n} \| Ax - b \|_2,$$

trong đó $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$với $m \geq n$. Bài toán này đã được trình bày và phân tích chi tiết trong Chương 1, bao gồm mô hình toán học, ý nghĩa hình học cũng như nghiệm bình phương tối thiểu.

Trong Chương 2, phân rã QR của ma trận đã được giới thiệu như một công cụ quan trọng trong đại số tuyến tính số. Cụ thể, một ma trận Acó thể được phân rã dưới dạng

$$A = QR,$$

trong đó Qlà ma trận có các cột trực giao chuẩn và Rlà ma trận tam giác trên. Các tính chất của phân rã QR, đặc biệt là tính bảo toàn chuẩn Euclid của ma trận trực giao, đã được trình bày ở chương trước.

Mục tiêu của chương này là vận dụng phân rã QR để giải bài toán bình phương tối thiểu đã nêu. Ý tưởng cơ bản của phương pháp là thay thế bài toán ban đầu bằng một bài toán tương đương nhưng đơn giản hơn về mặt tính toán, bằng cách khai thác tính trực giao của ma trận Q. Cách tiếp cận này cho phép giải bài toán bình phương tối thiểu mà không cần hình thành ma trận $A^T A$, từ đó cải thiện độ ổn định số so với phương pháp phương trình chuẩn.

Trên cơ sở đó, các mục tiếp theo sẽ trình bày chi tiết quá trình chuyển đổi bài toán bình phương tối thiểu bằng phân rã QR, quy trình giải cụ thể, minh họa bằng ví dụ số và so sánh với phương pháp phương trình chuẩn.

3.2. Chuyển đổi bài toán bình phương tối thiểu bằng phân rã QR

Xét bài toán bình phương tối thiểu

$$\min_{x \in \mathbb{R}^n} \| Ax - b \|_2,$$

trong đó $A \in \mathbb{R}^{m \times n}$với $m \geq n$. Theo phân rã QR đã trình bày ở Chương 2, ma trận Acó thể được viết dưới dạng

$$A = QR,$$

với $Q \in \mathbb{R}^{m \times n}$là ma trận có các cột trực giao chuẩn và $R \in \mathbb{R}^{n \times n}$là ma trận tam giác trên.

Thay phân rã này vào bài toán bình phương tối thiểu, ta có

$$\| Ax - b \|_2 = \| QRx - b \|_2.$$

Do các cột của Qtrực giao chuẩn, ma trận Qthỏa mãn tính chất bảo toàn chuẩn Euclid, tức là

$$\| QRx - b \|_2 = \| Q^T(QRx - b) \|_2.$$

Suy ra

$$\| Ax - b \|_2 = \| Rx - Q^T b \|_2.$$

Như vậy, bài toán bình phương tối thiểu ban đầu được chuyển đổi thành bài toán tương đương:

$$\min_{x \in \mathbb{R}^n} \| Rx - Q^T b \|_2.$$

Do Rlà ma trận tam giác trên kích thước $n \times n$, bài toán trên tương đương với việc giải hệ phương trình tam giác trên

$$Rx = Q^T b.$$

Trong trường hợp ma trận Acó hạng đầy đủ (full column rank), ma trận Rlà khả nghịch và hệ phương trình trên có nghiệm duy nhất. Khi đó, nghiệm thu được chính là nghiệm bình phương tối thiểu của bài toán ban đầu. Việc giải hệ tam giác trên có thể được thực hiện hiệu quả bằng phương pháp thế ngược, với chi phí tính toán thấp và độ ổn định số cao.

So với phương pháp phương trình chuẩn, cách tiếp cận dựa trên phân rã QR tránh được việc hình thành ma trận $A^T A$, vốn có thể làm gia tăng số điều kiện của bài toán. Nhờ đó, phương pháp QR được đánh giá là ổn định hơn về mặt số học và được khuyến nghị sử dụng trong các bài toán bình phương tối thiểu trong đại số tuyến tính số hiện đại.

3.3. Quy trình giải chi tiết

Dựa trên phép chuyển đổi bài toán bình phương tối thiểu bằng phân rã QR đã trình bày ở Mục 3.2, quy trình giải bài toán

$$\min_{x\in\mathbb{R}^n} \| Ax - b \|_2$$

có thể được thực hiện theo các bước sau.

Bước 1. Phân rã QR của ma trận A.

Thực hiện phân rã QR để biểu diễn ma trận Adưới dạng

$$A = QR,$$

trong đó $Q \in \mathbb{R}^{m\times n}$là ma trận có các cột trực giao chuẩn và $R \in \mathbb{R}^{n\times n}$là ma trận tam giác trên. Việc phân rã này có thể được thực hiện bằng các phương pháp đã trình bày trong Chương 2, chẳng hạn như Gram–Schmidt cải tiến, biến đổi Householder hoặc phép quay Givens.

Bước 2. Tính véc-tơ $\boldsymbol{Q^T b}$.

Sau khi thu được ma trận Q, tiến hành tính tích

$$c = Q^T b.$$

Do Qcó các cột trực giao chuẩn, phép nhân này có ý nghĩa như việc chiếu véc-tơ $b$lên không gian cột của ma trận A.

Bước 3. Giải hệ phương trình tam giác trên.

Giải hệ phương trình

$$Rx = c$$

để tìm nghiệm x. Vì Rlà ma trận tam giác trên và giả sử Acó hạng đầy đủ, hệ phương trình này có nghiệm duy nhất và có thể được giải hiệu quả bằng phương pháp thế ngược.

Bước 4. Thu được nghiệm bình phương tối thiểu.

Nghiệm $x$tìm được ở bước trên chính là nghiệm bình phương tối thiểu của bài toán ban đầu. Nghiệm này tối thiểu hóa chuẩn Euclid của sai số $\| Ax - b \|_2$.

Về mặt tính toán, quy trình trên có ưu điểm là tránh việc hình thành ma trận $A^T A$, do đó hạn chế sự khuếch đại sai số làm tròn. Đồng thời, việc giải hệ tam giác trên ở bước cuối có chi phí tính toán thấp và độ ổn định số cao. Chính vì vậy, phương pháp giải bài toán bình phương tối thiểu dựa trên phân rã QR được sử dụng rộng rãi trong thực hành tính toán số.

3.4. Minh hoạ bằng ví dụ số

Để minh họa cho phương pháp giải bài toán bình phương tối thiểu bằng phân rã QR, xét bài toán xấp xỉ dữ liệu sau:

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}.$$

Ma trận $A \in \mathbb{R}^{3\times 2}$có số hàng lớn hơn số cột, do đó hệ phương trình $Ax = b$là một hệ quá định và nói chung không có nghiệm chính xác.

Bước 1. Phân rã QR của ma trận A

Áp dụng phân rã QR (ví dụ bằng phương pháp Gram–Schmidt cải tiến), ta thu được:

$$Q = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \end{pmatrix}, R = \begin{pmatrix} \sqrt{3} & 2\sqrt{3} \\ 0 & \sqrt{2} \end{pmatrix}.$$

(Trong đó các cột của Q trực giao chuẩn và R là ma trận tam giác trên.)

Bước 2. Tính véc-tơ $Q^T b$

Ta có:

$$Q^T b = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{5}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Bước 3. Giải hệ phương trình tam giác trên

Giải hệ:

$$Rx = Q^T b$$

tức là

$$\begin{pmatrix} \sqrt{3} & 2\sqrt{3} \\ 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{5}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Từ phương trình thứ hai:

$$x_2 = \frac{1}{2}.$$

Thay vào phương trình thứ nhất:

$$\sqrt{3}x_1 + 2\sqrt{3} \cdot \frac{1}{2} = \frac{5}{\sqrt{3}} \Rightarrow x_1 = \frac{2}{3}.$$

Bước 4. Nghiệm bình phương tối thiểu

Nghiệm bình phương tối thiểu của bài toán là:

$$x = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{2} \end{pmatrix}.$$

Nghiệm này tương ứng với hàm xấp xỉ tuyến tính:

$$y = \frac{2}{3} + \frac{1}{2}x,$$

là đường thẳng "khớp tốt nhất" với dữ liệu theo nghĩa bình phương tối thiểu.

Nhận xét

Ví dụ trên cho thấy phương pháp phân rã QR cho phép giải bài toán bình phương tối thiểu một cách trực tiếp và có hệ thống. Toàn bộ quá trình chỉ yêu cầu phân rã QR và giải một hệ phương trình tam giác trên, tránh việc hình thành ma trận $A^T A$. Điều này góp phần nâng cao độ ổn định số của thuật toán so với phương pháp phương trình chuẩn.

3.5. So sánh phương pháp phân rã QR và phương pháp phương trình chuẩn

Trong chương này, bài toán bình phương tối thiểu đã được giải bằng phương pháp phân rã QR. Một cách tiếp cận cổ điển khác thường được trình bày trong các giáo trình cơ sở là phương pháp phương trình chuẩn (normal equations). Mục này nhằm so sánh hai phương pháp trên các khía cạnh quan trọng, từ đó làm rõ lý do lựa chọn phân rã QR trong thực hành tính toán số.

3.5.1. So sánh về mặt lý thuyết

Phương trình chuẩn xuất phát từ điều kiện cực tiểu của hàm sai số:

$$\min_x \ \| Ax - b \|_2^2,$$

dẫn đến hệ phương trình:

$$A^T Ax = A^T b.$$

Nếu A có hạng đầy đủ, ma trận $A^T A$ là xác định dương và hệ trên có nghiệm duy nhất.

Trong khi đó, phương pháp QR dựa trên việc phân rã:

$$A = QR,$$

với Q trực giao và R tam giác trên, từ đó bài toán LS được chuyển thành hệ:

$$Rx = Q^T b.$$

Về mặt lý thuyết, hai phương pháp cho cùng nghiệm bình phương tối thiểu trong số học chính xác. Tuy nhiên, sự khác biệt quan trọng xuất hiện khi xét đến tính toán số trên máy tính.

3.5.2. So sánh về độ ổn định số

Một trong những hạn chế lớn nhất của phương pháp phương trình chuẩn là việc hình thành ma trận $A^T A$. Số điều kiện của $A^T A$thỏa mãn:

$$\kappa(A^T A) = \kappa(A)^2,$$

nghĩa là sai số làm tròn có thể bị khuếch đại lên bình phương so với bài toán ban đầu.

Ngược lại, phương pháp QR chỉ sử dụng các phép biến đổi trực giao (Gram–Schmidt cải tiến, Householder, Givens), vốn bảo toàn chuẩn Euclid và không làm gia tăng đáng kể sai số làm tròn. Do đó, phương pháp QR được đánh giá là ổn định số hơn đáng kể so với phương trình chuẩn, đặc biệt khi ma trận Agần suy biến hoặc có số điều kiện lớn.

3.5.3. So sánh về độ phức tạp tính toán

Về mặt chi phí tính toán, hai phương pháp có sự khác biệt như sau:

- Phương trình chuẩn:
    - Tính $A^T A$: $O(mn^2)$
    - Giải hệ tuyến tính: $O(n^3)$
- Phân rã QR (Householder):
    - Phân rã QR: $O(2mn^2 - \frac{2}{3}n^3)$
    - Giải hệ tam giác: $O(n^2)$

Như vậy, phương trình chuẩn có chi phí tính toán thấp hơn đôi chút, đặc biệt khi mlớn và $n$nhỏ. Tuy nhiên, trong thực tế, sự chênh lệch này thường không đủ lớn để bù lại nhược điểm về độ ổn định số.

3.5.4. So sánh về khả năng mở rộng và ứng dụng thực tế

Trong các bài toán thực tế như hồi quy dữ liệu, xử lý tín hiệu hay học máy, ma trận A thường:

- có kích thước lớn,
- có thể gần suy biến,
- hoặc chứa nhiễu đo.

Trong những trường hợp này:

- Phương trình chuẩn dễ dẫn đến nghiệm kém chính xác hoặc không tin cậy.
- Phân rã QR (đặc biệt với Householder hoặc Givens) cho nghiệm ổn định và đáng tin cậy hơn.

Do đó, trong các thư viện số hiện đại như LAPACK, MATLAB, hay NumPy, phương pháp QR thường được ưu tiên hơn so với phương trình chuẩn khi giải bài toán bình phương tối thiểu.

3.5.5. Kết luận so sánh

Từ các phân tích và so sánh ở trên có thể thấy rằng phương pháp phương trình chuẩn có ưu điểm là cách tiếp cận trực tiếp và chi phí tính toán tương đối thấp. Tuy nhiên, do việc hình thành ma trận $A^T A$có thể làm gia tăng sai số làm tròn, phương pháp này kém ổn định về mặt số học, đặc biệt đối với các bài toán có ma trận hệ số gần suy biến hoặc có số điều kiện lớn.

Ngược lại, phương pháp phân rã QR dựa trên các phép biến đổi trực giao, giúp bảo toàn chuẩn và hạn chế sự khuếch đại sai số trong quá trình tính toán. Nhờ đó, phương pháp này cho nghiệm ổn định và đáng tin cậy hơn trong các ứng dụng thực tế. Vì những lý do trên, trong khuôn khổ báo cáo này, phương pháp phân rã QR được lựa chọn làm phương pháp chính để giải bài toán bình phương tối thiểu.

CHƯƠNG 4: CÀI ĐẶT VÀ THỰC NGHIỆM

TÀI LIỆU THAM KHẢO

[1] Matrix Computations 4th Edition by Gene H. Golub & Charles F. Van Loan
https://math.ecnu.edu.cn/~jypan/Teaching/books/2013%20Matrix%20Computations%204th.pdf