
15_Probabilistische_Lineare_Algebra

Martin Reißel

27. Juni 2022

Inhaltsverzeichnis

1 Probabilistische Lineare Algebra	1
1.1 Grundlagen	1
1.2 Probabilistische Verfahren	2
1.3 Fixed-Rank	4
1.4 Randomized SVD	5
1.5 Fixed-Precision	6
1.6 Zusammenfassung	7

1 Probabilistische Lineare Algebra

1.1 Grundlagen

Aus der linearen Algebra sind viele Matrix-Zerlegungen bekannt, z.B. LU, QR, SVD. Wir betrachten hier die SVD, also

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

mit U, V orthonormal und $\Sigma = \text{diag}_i(\sigma_i)$.

Hat A den Rang r , dann gilt

$$\begin{aligned}\sigma_i &> 0 & i = 1, \dots, r \\ \sigma_i &= 0 & i = r + 1, \dots\end{aligned}$$

und man erhält

$$\underbrace{A}_{\mathbb{R}^{m \times n}} = \underbrace{U_r}_{\mathbb{R}^{m \times r}} \underbrace{\Sigma_r}_{\mathbb{R}^{r \times r}} \underbrace{V_r^T}_{\mathbb{R}^{r \times n}}$$

mit $r \leq \min(m, n)$. Ist $r \ll \min(m, n)$, dann können Operationen mit A , wie z.B. Matrix-Vektor-Produkte, über die Zerlegung mit deutlich weniger Aufwand berechnet werden.

Die Spalten von U_r bilden eine Orthonormalbasis des Bildraums $R(A)$, d.h. sie enthalten die komplette Information über $R(A)$. Gilt nun

$$\sigma_1 \geq \dots \geq \sigma_k \geq \sigma_{k+1} \geq \dots \geq \sigma_r, \quad \sigma_1 \gg \sigma_{k+1}$$

dann hatten wir bei TSVD einfach $\sigma_{k+1} = \dots = \sigma_r = 0$ gesetzt und damit eine Matrix-Approximation durchgeführt

$$A = U\Sigma V^T \approx U_k \Sigma_k V_k^T.$$

Die Spalten von U_k sind eine Orthonormalbasis eines Unterraums von $R(A)$. Wegen $\sigma_1 \gg \sigma_{k+1}$ “hoffen” wir, das wir damit den “wesentlichen Teil” von $R(A)$ erwisch haben und damit $U_k \Sigma_k V_k^T$ eine gute Approximation von A ist.

Diese Vorgehensweise lässt sich verallgemeinern:

- für $A \in \mathbb{R}^{m \times n}$ sucht man eine Rang- k -Approximation

$$A \approx QC, \quad Q \in \mathbb{R}^{m \times k}, \quad C \in \mathbb{R}^{k \times n}$$

- dazu konstruiert man zunächst einen Unterraum von $R(A)$ der möglichst kleine Dimension k haben sollte, aber gleichzeitig die wesentlichen Abbildungseigenschaften von A wiedergeben soll
- eine Orthonormalbasis dieses Unterraums legen wir als Spalten in der Matrix

$$Q \in \mathbb{R}^{m \times k}$$

ab und approximieren A durch Orthogonalprojektion in diesen Unterraum, d.h.

$$A \approx QQ^T A = QC = B, \quad C = Q^T A \in \mathbb{R}^{k \times n}$$

- mit Q erzeugen wir also eine niedrig dimensionale Matrix $C = Q^T A$ so dass QC die Ausgangsmatrix A approximiert
- auf C wenden wir dann Standard-Zerlegungen (LU, QR, SVD) an, was aufgrund der niedrigen Dimension wenig Aufwand verursacht, und konstruieren daraus approximative Zerlegungen für die Ausgangsmatrix A

Wir gehen also zweistufig vor:

- bestimme Q
- berechne die Zerlegung von $C = Q^T A$ und darüber eine Zerlegung von $B = QC$, was dann eine approximative Zerlegung von A darstellt

Für den ersten Teil werden wir probabilistische Methoden benutzen.

Beispiel: Approximative SVD für $A \in \mathbb{R}^{m \times n}$

- $Q_k \in \mathbb{R}^{m \times k}$ orthonormal sei gegeben, $k \leq \min(m, n)$
- ist $C_k = Q_k^T A \in \mathbb{R}^{k \times n}$, dann ist

$$B_k = Q_k C_k = Q_k Q_k^T A \approx A$$

eine Rang- k -Approximation von A

- wir berechnen die SVD von C_k und erhalten

$$C_k = \tilde{U}_k \Sigma_k V_k^T, \quad \tilde{U}_k \in \mathbb{R}^{k \times k}, \quad \Sigma_k \in \mathbb{R}^{k \times k}, \quad V_k^T \in \mathbb{R}^{k \times n}$$

- mit $A \approx B_k = Q_k C_k$ folgt dann

$$A \approx Q_k C_k = \underbrace{Q_k \tilde{U}_k}_{U_k} \Sigma_k V_k^T,$$

mit $U_k \in \mathbb{R}^{m \times k}$ orthonormal

Die offene Frage ist jetzt, wie wir geeignete Matrizen Q_k finden können.

1.2 Probabilistische Verfahren

In diesem Abschnitt betrachten wir die Konstruktion von Q_k .

Ab jetzt lassen wir den Index k weg. Wir müssen folgendes Problem lösen:

- gegeben ist $A \in \mathbb{R}^{m \times n}$, $\varepsilon > 0$
- finde eine orthonormale Matrix $Q \in \mathbb{R}^{m \times k}$, $k = k(\varepsilon) \leq \min(m, n)$, mit

$$\|A - QQ^T A\| \leq \varepsilon,$$

d.h.

$$\|A - X\| \leq \varepsilon$$

mit $X = QQ^T A$ mit $\text{rang}(X) \leq k$

$k(\varepsilon)$ sollte so klein wie möglich sein. Als Matrix-Norm benutzen wir in der Regel

$$\|A\|_2 = \sqrt{\rho(A^T A)},$$

$$\|A\|_F = \sqrt{\text{tr}(A^T A)} = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2}.$$

Für orthonormal-invariante Matrix-Normen, d.h.

$$\|A\| = \|QA\| = \|A\tilde{Q}\| \quad \forall Q, \tilde{Q} \text{ orthonormal}$$

kennt man eine Optimal-Lösung dieses Problems.

Satz (Eckart, Young, Mirsky): Ist $A \in \mathbb{R}^{m \times n}$ mit SVD $A = U\Sigma V^T$ und die Matrixnorm $\|\cdot\|$ orthonormal invariant, dann gilt

$$\min_{X \in \mathbb{R}^{m \times n}, \text{rang}(X) \leq k} \|A - X\| = \sigma_{k+1}.$$

Das Minimum wird für

$$X_{\min} = U_k \Sigma_k V_k^T,$$

also TSVD, angenommen.

Die Normen $\|\cdot\|_2$ und $\|\cdot\|_F$ sind orthonormal invariant (Übung), somit können wir das Ergebnis hier in anwenden. In unserem Set-Up bedeutet das:

- in X_{\min} kommen nur die Spalten u_1, \dots, u_k von U vor, d.h.

$$Q = U_k = (u_1, \dots, u_k) \in \mathbb{R}^{m \times k}$$

- somit ist

$$\begin{aligned} QQ^T A &= U_k U_k^T U \Sigma V^T \\ &= U_k (I_k, 0) \Sigma V^T \\ &= U_k (\Sigma_k, 0) V^T \\ &= U_k \Sigma_k V_k^T \end{aligned}$$

und wir könnten wie folgt vorgehen:

- ▶ gebe $\varepsilon > 0$ vor
- ▶ suche k mit $\sigma_{k+1} < \varepsilon$
- ▶ benutze $Q = U_k$ und

$$B = QQ^T A = U_k \Sigma_k V_k^T$$

als Rang- k -Approximation

Das Problem dabei ist, dass man dazu eine komplette SVD von A benötigt.

Wir müssen also überlegen, wie wir "billiger" an Q kommen können. Dazu betrachten wir für $A \in \mathbb{R}^{m \times n}$ zwei Fälle:

- fixed-precision: gegeben $\varepsilon > 0$, suche Q orthonormal mit $\text{rang}(Q) \leq k(\varepsilon)$, wobei $k(\varepsilon)$ möglichst klein sein soll, und

$$\|A - QQ^T A\| \leq \varepsilon$$

- fixed-rank: k, p ist gegeben, suche $Q \in \mathbb{R}^{m \times (k+p)}$ orthonormal mit $\text{rang}(Q) \leq k$ und

$$\|A - QQ^T A\| \approx \min_{X \in \mathbb{R}^{m \times n}, \text{rang}(X) \leq k} \|A - X\| = \sigma_{k+1}$$

Bei fixed-rank lassen wir bei Q mehr Spalten zu als für eine Rang- k -Approximation unbedingt nötig wären (p). Dieser Freiheitsgrad wird später eine wichtige Rolle bei der Genauigkeit probabilistischer Methoden spielen.

Wir untersuchen zunächst fixed-rank und knacken damit dann später auch fixed-precision.

1.3 Fixed-Rank

Wie kann der “Zufall” bei der Konstruktion von Q helfen?

Wir betrachten dazu zunächst $A \in \mathbb{R}^{m \times n}$ mit $\text{rang}(A) = k$:

- ziehe k Zufallsvektoren $\omega^{(i)}$ (Verteilung zunächst uninteressant) und berechne

$$y^{(i)} = A\omega^{(i)}, \quad i = 1, \dots, k$$

- mit hoher Wahrscheinlichkeit gilt:
 - ▶ die $\omega^{(i)}$ sind linear unabhängig
 - ▶ $\omega^{(i)} \notin N(A)$
- damit sind dann auch die $y^{(i)}$ mit hoher Wahrscheinlichkeit linear unabhängig
- wegen $\text{rang}(A) = k$ gilt dann

$$R(A) = \text{span}(y^{(1)}, \dots, y^{(k)}),$$

d.h. $Y = (y^{(1)}, \dots, y^{(k)})$ spannt $R(A)$ auf

- orthonormalisieren wir jetzt Y (modifizierter Gram-Schmidt, QR, etc.), so erhalten wir $Q \in \mathbb{R}^{m \times k}$

Jetzt sei $A = A_0 + E$, $\text{rang}(A_0) = k$, wobei E sei eine “kleine” Störung ist.

- wir suchen einen Unterraum in $R(A)$, der einen möglichst “großen Teil” von $R(A_0)$ abdeckt, aber nicht unbedingt die kleinstmögliche Dimension k hat sondern Dimension $k + p$
- wir betrachten wieder

$$y^{(i)} = A\omega^{(i)} = A_0\omega^{(i)} + E\omega^{(i)}$$

- der “Störanteil” $E\omega^{(i)}$ “schiebt” $A_0\omega^{(i)}$ eventuell aus $R(A_0)$ hinaus
- wenn wir genau k verschiedene $y^{(i)}$ benutzen, dann ist die Wahrscheinlichkeit hoch, dass wir “wichtige Teile” von $R(A_0)$ nicht erfassen
- deshalb benutzen wir $k + p$ Vektoren $y^{(i)}$, $p \in \mathbb{N}$

Fassen wir diese Schritte zusammen, so erhalten wir das **Fixed-Rank-Verfahren**:

- gegeben ist $A \in \mathbb{R}^{m \times n}$, $k, p \in \mathbb{N}$
- bestimme eine “zufällige” Matrix $\Omega \in \mathbb{R}^{n \times (k+p)}$
- berechne damit

$$Y = A\Omega \in \mathbb{R}^{m \times (k+p)}$$

- orthonormiere die Spalten von Y und erzeuge damit die orthonormale Matrix

$$Q \in \mathbb{R}^{m \times (k+p)}$$

Dabei sind folgende Fragen noch offen:

- welches Ω und welches p soll benutzt werden?
- wie orthonormalisiert man Y , das in der Regel schlechte Kondition hat (Spalten sind “fast” linear abhängig)?
- wie ist der Aufwand, die Genauigkeit?
- wie geht fixed-precision?

Satz: $A \in \mathbb{R}^{m \times n}$, $k \geq 2$, $p \geq 2$, $k + p \leq \min(m, n)$,

$$\Omega = (\omega_{ij})_{i,j} \in \mathbb{R}^{n \times (k+p)}, \quad \omega_{ij} \sim \mathcal{N}(0, 1) \text{ iid},$$

dann gilt

$$\mathbb{E}_\Omega(\|A - QQ^T A\|_2) \leq (1 + \delta)\sigma_{k+1},$$

$$\delta = \frac{4\sqrt{k+p}}{p-1} \sqrt{\min(m, n)}$$

bzw.

$$\mathbb{P}(\|A - QQ^T A\|_2 \leq (1 + \tilde{\delta})\sigma_{k+1}) \geq 1 - 3p^{-p},$$

$$\tilde{\delta} = 9\sqrt{k+p}\sqrt{\min(m, n)}.$$

Bemerkung:

- $\min_{X \in \mathbb{R}^{m \times n}, \text{rang}(X) \leq k} \|A - X\|_2 = \sigma_{k+1}$ so dass $\delta, \tilde{\delta}$ die Abweichung vom Optimum angibt
- $\delta, \tilde{\delta}$ wächst nur schwach in k, m, n
- es ist (unabhängig von k, m, n)

$$3p^{-p} \approx \begin{cases} 10^{-3} & p = 5 \\ 3 \cdot 10^{-10} & p = 10 \end{cases},$$

so dass für recht kleine p -Werte $\mathbb{P}(\|A - QQ^T A\|_2 \leq (1 + \tilde{\delta})\sigma_{k+1}) \approx 1$

- fallen die Singulärwerte schnell ab, so dass σ_{k+1} klein ist, dann sind die Approximationen sehr gut

1.4 Randomized SVD

Wir wenden jetzt die Methoden des letzten Abschnitts, um eine approximative SVD von A zu berechnen.

Eine direkte Anwendung liefert oft schlechte Ergebnisse, insbesondere wenn die Singulärwerte von A nur langsam abfallen. Statt $Y = A\Omega$ benutzt man dann die Power-Methode

$$Y = (AA^T)^q A\Omega,$$

wobei in der Praxis meist $q = 1$ oder $q = 2$ ausreicht.

Es gilt $R((AA^T)^q A) \subset R(A)$ und mit $A = U\Sigma V^T$ folgt

$$AA^T = U\Sigma\Sigma^T U^T$$

also

$$\begin{aligned} (AA^T)^q A &= (U\Sigma\Sigma^T U^T)^q U\Sigma V^T \\ &= U(\Sigma\Sigma^T)^q \Sigma V^T \\ &= U\Sigma_{q+1} V^T \end{aligned}$$

mit

$$\Sigma_{q+1} = \text{diag}(\sigma_i^{q+1}),$$

wobei σ_i^{q+1} für $q > 0$ schneller abfällt als σ_i .

Damit erhalten wir schließlich folgenden Algorithmus für RSVD:

- $A \in \mathbb{R}^{m \times n}$, $k \in \mathbb{N}$, $p, q \in \mathbb{N}_0$
- Schritt 1:

- ▶ $\Omega = (\omega_{ij})_{i,j} \in \mathbb{R}^{n \times (k+p)}$, $\omega_{ij} \sim \mathcal{N}(0, 1)$ iid
- ▶ $Y = (AA^T)^q A\Omega \in \mathbb{R}^{m \times (k+p)}$
- ▶ orthonormalisiere die Spalten von Y und lege das Ergebnis in $Q \in \mathbb{R}^{m \times (k+p)}$ ab
- Schritt 2:
 - ▶ $C = Q^T A \in \mathbb{R}^{(k+p) \times n}$
 - ▶ berechne SVD von C , $C = \tilde{U}\Sigma V^T$
 - ▶ setze $U = Q\tilde{U}$
- insgesamt ist dann $A \approx U\Sigma V^T$

Bemerkung:

- man berechnet eine Rang- $(k+p)$ -Approximation um eine (gute) Näherung für die ersten k Singulärwerte und -vektoren zu erhalten
- die Produkte $AA^T A\Omega$ sind anfällig gegen Rundungsfehler, weshalb man nach jedem Schritt orthonormalisieren sollte

Satz: $A \in \mathbb{R}^{m \times n}$, $q \in \mathbb{N}_0$, $2 \leq k \leq \frac{1}{2} \min(m, n)$, dann gilt für U, Σ, V aus der RSVD mit $p = k$

$$\mathbb{E}_\Omega(\|A - U\Sigma V^T\|_2) \leq (1 + \delta)^{\frac{1}{2q+1}} \sigma_{k+1},$$

$$\delta = 4\sqrt{\frac{2 \min(m, n)}{k-1}}$$

Da wir nur an den ersten k Approximationen der Singulärwerte und -vektoren interessiert sind, ist es naheliegend, die σ_i für $i > k$ abzuschneiden. Wir ersetzen oben Σ durch

$$\Sigma_k = \text{diag}(\sigma_1, \dots, \sigma_k, 0, \dots, 0),$$

d.h. wir benutzen eine TSVD von C in Schritt 2. In diesem Fall ändert sich die Fehlerabschätzung des letzten Satzes geringfügig.

Satz: $A \in \mathbb{R}^{m \times n}$, $q \in \mathbb{N}_0$, $2 \leq k \leq \frac{1}{2} \min(m, n)$, dann gilt für U, Σ_k, V aus der RSVD mit $p = k$

$$\mathbb{E}_\Omega(\|A - U\Sigma_k V^T\|_2) \leq (1 + \delta)^{\frac{1}{2q+1}} \sigma_{k+1} + \sigma_{k+1},$$

$$\delta = 4\sqrt{\frac{2 \min(m, n)}{k-1}}$$

Damit ist die fixed-rank Variante der RSVD soweit erledigt.

1.5 Fixed-Precision

Das fixed-precision Problem könnte wie folgt auf die fixed-rank Variante zurückgeführt werden:

- starte mit fixed-rank mit kleinem k
- teste, ob $\|A - QQ^T A\|_2 \leq \varepsilon$
- wenn nicht, dann erhöhe k und wiederhole den vorherigen Schritt

Das Berechnen der Matrixnorm $\|A - QQ^T A\|_2$ ist in der Regel zu aufwendig, deshalb benötigt man dafür einen einfachen a-posteriori Fehlerindikator. Auch hier hilft uns wieder der “Zufall” weiter.

Lemma: Sei $B \in \mathbb{R}^{m \times n}$, $\omega^{(j)} \in \mathbb{R}^n$, $j = 1, \dots, r$, $\omega_i^{(j)} \sim \mathcal{N}(0, 1)$ iid und $\alpha > 1$. Dann gilt

$$\mathbb{P}(\|B\|_2 \leq \alpha \sqrt{\frac{2}{\pi}} \max_{j=1, \dots, r} \|B\omega^{(j)}\|_2) \geq 1 - \alpha^{-r}.$$

Für die Praxis bedeutet das:

- erzeuge $\omega^{(j)}$, $j = 1, \dots, r$ für r klein
- mit $m_r = \max_{j=1, \dots, r} \|(A - QQ^T A)\omega^{(j)}\|_2$ und $\alpha = 10$ folgt

$$\mathbb{P}(\|A - QQ^T A\|_2 \leq 10\sqrt{\frac{2}{\pi}} m_r) \geq 1 - 10^{-r},$$

so dass $r \leq 10$ ausreichend ist

Prinzipiell wenden wir beim diesem Fehlerschätzer die gleichen Operationen wie bei der Gaussian-Projection

$$Y = A\Omega, \quad \Omega = (\omega_{ij})_{i,j} \in \mathbb{R}^{n \times (k+p)}, \quad \omega_{ij} \sim \mathcal{N}(0, 1) \text{ iid}$$

an, nur dass wir statt der Matrix Ω einzelne Vektoren $\omega^{(j)}$ benutzen.

Ordnen wir die Berechnungsschritte etwas um, dann können wir den Fehlerschätzer praktisch ohne Mehraufwand in unser fixed-rank-Verfahren integrieren:

- gegeben sind $A \in \mathbb{R}^{m \times n}$, $r \in \mathbb{N}$, $\varepsilon > 0$:
 - ▶ erzeuge $\omega^{(j)} \in \mathbb{R}^n$, $j = 1, \dots, r$, $\omega_i^{(j)} \sim \mathcal{N}(0, 1)$ iid
 - ▶ $y^{(j)} = A\omega^{(j)}$, $j = 1, \dots, r$
 - ▶ $k = 0$, $Q^{(0)} = [\]$ (leer)
 - ▶ wiederhole bis $\max_{j=1, \dots, r} \|y^{(k+j)}\|_2 \leq \frac{\varepsilon}{10\sqrt{\frac{2}{\pi}}}$:
 - $k := k + 1$
 - $y^{(k)} := (I - Q^{(k-1)}Q^{(k-1)T})y^{(k)}$
 - $q^{(k)} := \frac{y^{(k)}}{\|y^{(k)}\|_2}$
 - $Q^{(k)} = [Q^{(k-1)}, q^{(k)}]$
 - erzeuge $\omega^{(k+r)} \in \mathbb{R}^n$, $\omega^{(k+r)} \sim \mathcal{N}(0, 1)$ iid
 - $y^{(k+r)} := (I - Q^{(k)}Q^{(k)T})A\omega^{(k+r)}$
 - $y^{(l)} := y^{(l)} - q^{(k)T}y^{(l)}q^{(k)}$ für $l = k + 1, k + 2, \dots, k + r - 1$
 - ▶ mit $Q = Q^{(j)}$ gilt dann

$$\mathbb{P}(\|A - QQ^T A\|_2 \leq \varepsilon) \geq 1 - 10^{-r} \min(m, n)$$

Damit ist Q bestimmt. Der Rest der RSVD ist identisch zum fixed-rank-Fall, insbesondere kann dabei auch wieder die Power-Methode benutzt werden.

1.6 Zusammenfassung

Wir haben probabilistische Methoden für approximative Berechnung von TSVD betrachtet und Algorithmen für fixed-rank und fixed-precision Aufgabenstellungen vorgestellt.