Conjugate Gradient Method

Introduction

Gram-Schmidt

$$\boldsymbol{R} = \boldsymbol{r}_i \otimes \boldsymbol{e}_i$$

$$\boldsymbol{R}^T \cdot \boldsymbol{R} = \boldsymbol{I}$$

Foundation of Conjugate Gradient Method

$(\boldsymbol{P}, \boldsymbol{P}^{-T})$ are a pair of covariant & contravariant transformations, and $(\boldsymbol{p}_i, \boldsymbol{p}^i)$ are the vectors that represent:

$$\begin{cases} \boldsymbol{P} = \boldsymbol{p}_i \otimes \boldsymbol{e}_i \\ \boldsymbol{P}^{-T} = \boldsymbol{p}^i \otimes \boldsymbol{e}_i \\ \boldsymbol{p}^i \cdot \boldsymbol{p}_j = \delta_{ij} \end{cases}$$

The transformation that transforms covariant space to contravariant is:

$$\boldsymbol{\Phi} = \boldsymbol{p}^i \otimes \boldsymbol{p}^i = \boldsymbol{P}^{-T} \cdot \boldsymbol{P}^{-1}$$

This transformation $\boldsymbol{\Phi}$ is positive semi-definite.

$\boldsymbol{Lemma\ 1}$: for any covariant transformation $\boldsymbol{P}$, the covariant-to-contravariant transformation $\boldsymbol{\Phi} = \boldsymbol{P}^{-T} \cdot \boldsymbol{P}^{-1}$ is also the covariant-to-contravariant transformation for $\boldsymbol{Q} = \boldsymbol{P} \cdot \boldsymbol{R}$.

$\boldsymbol{Proof\ 1}$: If the covariant transformation happens after a unitary transformation, then the covariant-to-contravariant transformation is independent of the unitary transformation:

$$\boldsymbol{Q} = \boldsymbol{P} \cdot \boldsymbol{R}$$

$$\boldsymbol{\Psi} = \boldsymbol{Q}^{-T} \cdot \boldsymbol{Q}^{-1} = \boldsymbol{P}^{-T} \cdot \boldsymbol{R}^{-T} \cdot \boldsymbol{R}^{-1} \cdot \boldsymbol{P}^{-1} = \boldsymbol{\Phi} = \boldsymbol{P}^{-T} \cdot \boldsymbol{P}^{-1}$$

$\boldsymbol{Lemma\ 2}$: any positive semi-definite tensor can be represented as a tensor multiply (dot product) by its transpose.

$\boldsymbol{Proof\ 2}$: $\boldsymbol{A}$ is positive semi-definite tensor, by spectrum decomposition:

$$\boldsymbol{A} = A_{ij} \boldsymbol{e}_i \otimes \boldsymbol{e}_j = \sum_i \lambda_i \boldsymbol{n}_i \otimes \boldsymbol{n}_i \quad where\ A_{ij} = A_{ji}, \lambda_i \geq 0$$

$$\boldsymbol{A} = \boldsymbol{B}^T \cdot \boldsymbol{B} \quad where\ \boldsymbol{B} = \boldsymbol{A}^{\frac{1}{2}} = \sum_i \lambda_i^{\frac{1}{2}} \boldsymbol{n}_i \otimes \boldsymbol{n}_i$$

Now if we claim $\boldsymbol{A} = \boldsymbol{P}^{-T} \cdot \boldsymbol{P}^{-1}$, by the same procedure we find:

$$\boldsymbol{P} = \boldsymbol{A}^{-\frac{1}{2}} = \lambda_i^{-\frac{1}{2}} \boldsymbol{n}_i \otimes \boldsymbol{n}_i$$

Then with Lemma 1:

$$\boldsymbol{A} = \boldsymbol{Q}^{-T} \cdot \boldsymbol{Q}^{-1} \quad where\ \boldsymbol{Q} = \boldsymbol{P} \cdot \boldsymbol{R}$$

$$det(\boldsymbol{A}) = det(\boldsymbol{Q})$$

This has guarantee that, by selecting any initial directions within $\boldsymbol{A}$ space, we can always generate a series of conjugate vectors with respect to $\boldsymbol{A}$ ($\boldsymbol{A - orthonormal}$ basis) that span the entire $\boldsymbol{A}$ space through the Generalized Gram-Schmidt Process.

Implementation of Conjugate Gradient Method

Relationship with Krylov Subspace