Preparazione alla Discussione Homework
Calcolo Numerico

Homework 1: Zeri di Funzione

Focus: Risoluzione di equazioni non lineari $f(x) = 0$ tramite metodi iterativi.

Teoremi e Concetti Teorici da Sapere

- **Teorema di Bolzano (degli Zeri):** Fondamentale per il metodo di bisezione. Garantisce l'esistenza di uno zero in un intervallo $[a, b]$ se f è continua e $f(a)f(b) < 0$.
- **Teorema di Convergenza del Metodo di Bisezione:** L'errore si dimezza ad ogni passo; la convergenza è lineare ma garantita.
- **Teorema del Punto Fisso (Contrazione):** Per il metodo $x_{k+1} = g(x_k)$. Condizione sufficiente per la convergenza locale è che $|g'(x)| < 1$ in un intorno della soluzione.
- **Teorema di Convergenza di Newton:** Convergenza quadratica (molto rapida) subordinata alla scelta del punto iniziale (sufficientemente vicino alla soluzione) e alla condizione $f'(x) \neq 0$.

Traccia del Discorso

In questo primo homework ho analizzato il problema della ricerca degli zeri per diverse funzioni non lineari, come polinomi, logaritmi ed esponenziali. Ho confrontato tre metodi: Bisezione, Iterazione di Punto Fisso e Newton.

Dalle analisi grafiche preliminari ho isolato gli intervalli di confidenza. Ho osservato che la **Bisezione** è molto robusta: converge sempre purché l'intervallo iniziale racchiuda lo zero, ma è lenta.

Per il **Punto Fisso**, la scelta della funzione $g(x)$ è stata critica: ho verificato sperimentalmente che il metodo converge solo quando la derivata prima della funzione di iterazione è in modulo minore di 1 vicino alla soluzione.

Infine, il **Metodo di Newton** si è dimostrato il più efficiente in termini di numero di iterazioni grazie alla sua convergenza quadratica, ma ho notato che è sensibile alla scelta del punto iniziale: se scelto male, il metodo può divergere o oscillare. Ho anche analizzato come la tolleranza scelta influenzi l'accuratezza finale e il costo computazionale.

Homework 2: Metodo del Gradiente (Ottimizzazione)

Focus: Minimizzazione di funzioni $f : \mathbb{R}^n \to \mathbb{R}$ e applicazione ai minimi quadrati.

Teoremi e Concetti Teorici da Sapere

- **Teorema di Fermat (punti stazionari):** $\nabla f(x) = 0$ è condizione necessaria per i minimi (interni).
- **Direzione di Discesa:** La direzione del gradiente negativo $-\nabla f(x)$ garantisce la decrescita locale della funzione.
- **Condizioni di Armijo (Backtracking):** Necessarie per garantire la convergenza quando si usa un passo variabile, evitando passi troppo lunghi o troppo corti.
- **Teorema di convergenza del Gradiente:** Sotto ipotesi di convessità e lipschitzianità del gradiente.
- **Minimi Locali vs Globali:** Distinzione cruciale per funzioni non convesse come la funzione di Rosenbrock o funzioni multimodali.

Traccia del Discorso

Nel secondo homework ho implementato il metodo di discesa del Gradiente per funzioni in più variabili. L'obiettivo era trovare i minimi di funzioni complesse, inclusa la famosa "Rosenbrock valley" e funzioni multimodali.

Ho implementato due strategie per il passo (learning rate): **passo fisso** e **backtracking**. Ho osservato che il passo fisso è rischioso: se troppo grande il metodo diverge (zigzag), se troppo piccolo è lentissimo. Il backtracking risolve questo problema adattando il passo dinamicamente.

Un punto interessante è stato l'analisi dei **Minimi Quadrati**: ho visto come il problema di regressione possa essere formulato come un problema di ottimizzazione. Infine, ho introdotto la **Regolarizzazione** (Ridge) e testato il **Gradiente Stocastico (SGD)**, notando come l'uso di mini-batch introduca "rumore" nel percorso di discesa ma permetta di scappare da minimi locali superficiali e sia computazionalmente più efficiente su grandi dataset.

Homework 3: Analisi Dati (Regressione)

Focus: Approssimazione di dati tramite modelli polinomiali e regressione lineare multipla.

Teoremi e Concetti Teorici da Sapere

- **Teorema di Esistenza e Unicità dei Minimi Quadrati:** La soluzione alle equazioni normali $A^T Ax = A^T b$.
- **Decomposizione SVD (Singular Value Decomposition):** Essenziale per risolvere il problema dei minimi quadrati in modo numericamente stabile, specialmente con matrici mal condizionate come quella di Vandermonde.

- **Fenomeno di Runge:** Legato all'uso di polinomi di grado elevato su nodi equispaziati, che causa oscillazioni agli estremi dell'intervallo.
- **Bias-Variance Trade-off:** Concetto teorico alla base dell'underfitting (modello troppo semplice) e dell'overfitting (modello che segue il rumore).

Traccia del Discorso

Il terzo homework si concentra sull'adattamento dei dati (data fitting). Ho affrontato il problema della **Regressione Polinomiale** risolvendolo come un problema di minimi quadrati lineari $\min \|X\alpha - y\|_2^2$.

Per la risoluzione numerica ho utilizzato la **SVD** invece delle equazioni normali classiche, poiché la matrice di Vandermonde tende a essere molto mal condizionata al crescere del grado del polinomio.

Ho sperimentato il fenomeno dell'**Overfitting**: aumentando troppo il grado del polinomio, l'errore sui dati di training scende, ma il modello apprende il "rumore" invece del trend, oscillando violentemente. Viceversa, un grado troppo basso porta all'Underfitting. Ho concluso applicando questi concetti a un dataset reale scaricato da Kaggle, utilizzando anche metriche come l'MSE e R^2 per validare la qualità del modello.

Homework 4: Imaging (SVD & Problemi Inversi)

Focus: Compressione immagini e ricostruzione (deblurring).

Teoremi e Concetti Teorici da Sapere

- **Teorema di Eckart-Young-Mirsky:** Afferma che la miglior approssimazione di rango p di una matrice A (in norma 2 o Frobenius) si ottiene troncando la sua SVD ai primi p valori singolari.
- **Problemi Mal Posti (Hadamard):** Il deblurring è un problema inverso dove la soluzione è instabile rispetto al rumore (piccole variazioni nei dati causano enormi variazioni nella soluzione).
- **Regolarizzazione di Tikhonov:** Sostituzione del problema originale con uno che minimizza $\|Ax - b\|^2 + \lambda\|x\|^2$ per imporre stabilità.
- **Principio di Massima Discrepanza:** Metodo per scegliere il parametro di regolarizzazione λ ottimale in base al livello di rumore.

Traccia del Discorso

L'ultimo homework applica l'algebra lineare al processamento delle immagini. Nella prima parte, ho utilizzato la **SVD** per la compressione. Troncando la somma ai primi p valori singolari, ho ottenuto un'approssimazione che conserva l'informazione principale (basso rango) scartando i dettagli fini meno energetici. Ho analizzato il compromesso tra fattore di compressione ed errore relativo.

Nella seconda parte, ho affrontato il **Deblurring**, un problema inverso mal posto. Ho verificato che la "soluzione naive" (inversione diretta o minimi quadrati semplici) è inutilizzabile perché amplifica esponenzialmente il rumore presente nell'immagine.

Per risolvere questo problema, ho applicato tecniche di **Regolarizzazione**:

1. **Tikhonov**, che impone una penalità sulla norma della soluzione (energia).

2. **Total Variation**, che preserva meglio i bordi dell'immagine.

Ho utilizzato metodi iterativi come **CGLS** e scelto il parametro di regolarizzazione ottimale tramite il principio di Massima Discrepanza, bilanciando la fedeltà ai dati con la regolarità dell'immagine ricostruita.