

Numerical optimization, Problem sheet 9

1. Justify that if f is strongly convex ($mI \leq f''$ for some $m > 0$), then in notation from the notes on quasi-Newton method $\langle p_i, q_i \rangle > 0$. Show by example that this may fail for nonconvex f .
2. Check that update operator in Broyden family maps plane P spanned by $S_i q_i$ and p_i into itself and maps orthogonal complement of P to 0.
3. Perform calculation checking that Sherman-Morrison-Woodbury formula really gives inverse.
4. Verify that BFGS formula given in part about LBFGS agrees with BFGS formula for S_i given earlier.
5. Rewrite momentum formula $y_{i+1} = y_i - \eta_i A_i + \theta_i (y_i - y_{i-1})$ written in the notes in terms of A_i as matrix transforming pair $(y_i, \gamma y_{i-1})$ into $(y_{i+1}, \gamma y_i)$. Show by example that this matrix may have norm bigger than γ .
Hint: It is enough to look at A_i in diagonal form. Then the matrix above consists of 2 by 2 blocks and it is enough to compute norm of single block.
Remark: This is the reason for linear factor in estimate for convergence rate of momentum for quadratic functions.
Remark: More involved calculation proves that all eigenvalues of matrix above have absolute value γ .