

szepesvari 本: 数式補足

2.4 Dynamic programming algorithms for solving MDPs

Q^k に対してグリーディな方策に関するバウンドの証明

$$V^\pi(x) \geq V^*(x) - \frac{2}{1-\gamma} \|Q - Q^*\|_\infty \quad (1)$$