Key Theorems in Optimization

Unconstrained Optimization

Existence

P1 If f cont. on compact S, then $\min_S f$ attained.
P2 If $L(\alpha) = \{f \le \alpha\}$ is compact, $\exists$ global minimizer in $L(\alpha)$.
P3 If $f(x) \to +\infty$ as $\|x\| \to \infty$, then all level sets compact.

OPT Conditions

FONC $\nabla f(x^*) = 0$.
SONC $\nabla f(x^*) = 0,\ \nabla^2 f(x^*) \succeq 0$.
SOSC $\nabla f = 0,\ \nabla^2 f(x^*) \succ 0 \implies$ strict local min.

Iterative Methods

Armijo $f(x+\alpha d) \le f(x) + \gamma\alpha\nabla f^T d$ (backtracking).
GD $x_{k+1} = x_k - \alpha_k \nabla f$, $\nabla f(x_k) \to 0$.
Newton $x_{k+1} = x_k - [\nabla^2 f]^{-1}\nabla f$, quad. conv. near sol.
CG For $f = \frac{1}{2}x^T Q x + c^T x$, CG finds minimizer in $\le n$ steps.

Constrained Optimization

KKT Conditions

stationarity: $\nabla_x L = 0$
primal: $g = 0,\ h \le 0$
dual: $\rho \ge 0$
comp. slack: $\rho^T h = 0$
$L = f + \lambda^T g + \rho^T h.$

Constrained Methods

Penalty $F_\epsilon = f + \dfrac{1}{2\epsilon}\sum g_i^2 + \dfrac{1}{2\epsilon}\sum [h_j^+]^2$, $\epsilon \to 0 \implies$ feasible sol.
AugL $\mathcal{L}_\epsilon = f + \lambda^T g + \rho^T h^+ + \dfrac{1}{2\epsilon}(\|g\|^2 + \|h^+\|^2)$,
$x_{k+1} = \arg\min_x \mathcal{L}$, $\lambda_{k+1} = \lambda + g/\epsilon$, $\rho_{k+1} = \max\{0, \rho + h/\epsilon\}$.
SQP QP: $\min_d \nabla f^T d + \frac{1}{2}d^T B d$ s.t. linearized cons.;
$B \approx \nabla^2_{xx} L$, $\alpha = 1 \Rightarrow$ superlinear conv.
RQP Solve $\left[\begin{smallmatrix} H & J^T \\ J & 0 \end{smallmatrix}\right](d, \delta) = -(\nabla f, g)$,
$x_{k+1} = x_k + d$, $\lambda_{k+1} = \lambda + \delta$, superlinear conv.

Additional Exam Topics

Constraint Qualifications

LICQ	Active gradients lin. indep.
MFCQ	LICQ $+\ \exists d$ s.t. $\nabla g^T d = 0, \nabla h_a^T d < 0$.
Slater	$\exists$ strict feasible $\hat{x}$: $g(\hat{x}) = 0,\ h(\hat{x}) < 0$.

Duality

Dual func: $q(\lambda, \rho) = \inf_x L(x, \lambda, \rho)$, dual prob: $\max_{\rho \ge 0} q$. Weak: $f(x) \ge q$. Strong: convex+Slater $\Rightarrow$ zero gap.

Line Search & Trust Region

Wolfe: $f(x+\alpha d) \le f + c_1\alpha\nabla f^T d$, $\nabla f(x+\alpha d)^T d \ge c_2 \nabla f^T d$.
Trust-Region: $\min_{\|d\| \le \Delta} m(d) = \nabla f^T d + \frac{1}{2}d^T B d$, Cauchy point $d_C = -\dfrac{\nabla f^T \nabla f}{\nabla f^T B \nabla f}\nabla f$.

Convergence Rates

- GD (L-smooth): $f(x_k) - f^* \le \frac{L\|x_0 - x^*\|^2}{2k}$.
- GD (μ-strong convex): $\|x_k - x^*\| \le (1 - \mu/L)^k \|x_0 - x^*\|$.
- Newton: quadratic local rate.
- SQP/RQP: superlinear (quadratic if exact Hessian).

Barrier/IP Methods

Solve $\min_x f(x) - \mu\sum_{j=1}^p \ln(-h_j(x))$ s.t. $g = 0$, $\mu \downarrow 0$. Perturbed KKT $\mu/h_j = \rho_j$.

Quasi-Newton Updates

$B_{k+1}^{DFP} = B_k + \frac{s_k s_k^T}{s_k^T y_k} - \frac{B_k y_k y_k^T B_k}{y_k^T B_k y_k}$, $B_{k+1}^{BFGS} = B_k - \frac{B_k y_k y_k^T B_k}{y_k^T B_k y_k} + \frac{s_k s_k^T}{s_k^T y_k}$.

Special and Heuristic Methods

Levenberg–Marquardt	For scalar $f(x) = 0$, with $\bar{x} = x_k - \frac{f(x_k)}{f'(x_k)}$: $x_{k+1} = x_k - \frac{2f(x_k)}{f'(x_k) + f'(\bar{x})}$.
Coordinate Descent	Optimize one coordinate at a time via exact line search: cycle $i = 1, \dots, n$.
False Position	Approximate Newton: $x_{k+1} = x_k - f'(x_k)/\big((f'(x_{k-1}) - f'(x_k))/(x_{k-1} - x_k)\big)$.
Pareto Optimality	For $f: R^n \to R^s$, x^* Pareto if no feasible $\tilde{x}$ has $f_i(\tilde{x}) \le f_i(x^*)\ \forall i$ and $<$ for some i.

$B_{k+1}^{DFP} = B_k + \frac{s_k s_k^T}{s_k^T y_k} - \frac{B_k y_k y_k^T B_k}{y_k^T B_k y_k}$, $B_{k+1}^{BFGS} = B_k - \frac{B_k y_k y_k^T B_k}{y_k^T B_k y_k} + \frac{s_k s_k^T}{s_k^T y_k}$.