

Index

- ε -optimal and feasible solution, 241
- ε -optimal solution, 206

- absolutely permutation symmetric function, 181
- absolutely symmetric function, 179
- accelerated dual block proximal gradient method, 373
- AD-LPMM, *see* alternating direction linearized prox method of multipliers
- AD-PMM, *see* alternating direction proximal method of multipliers
- ADBPG, *see* accelerated dual block proximal gradient
- adjoint transformation, 11
- ADMM, *see* alternating direction method of multipliers
- affine hull, 3
- affine set, 3
- alternating direction linearized prox method of multipliers, 426
- alternating direction method of multipliers, 425
- alternating direction proximal method of multipliers, 425
- alternating minimization, 405
- alternating projection method, 211
- augmented Lagrangian, 425
- augmented Lagrangian method, 423

- ball-pen function, 99, 125
- basis, 2
- biconjugate function, 89
- bidual space, 10
- block descent lemma, 336
- block Lipschitz constant, 333
- block proximal gradient method, 338
- block sufficient decrease lemma, 337
- box, 5

- Bregman distance, 248

- Cartesian product, 7
- CBPG, *see* cyclic block proximal gradient method
- chain rule, 59
- closed ball, 2
- closed function, 14
- closed line segment, 3
- coercive, 20
- compact set, 20, 42
- complexity, 206
- composite model, 78
- conditional gradient method, 379
- conditional gradient norm, 381
- cone, 27
- conjugate correspondence theorem, 123
- conjugate function, 87
- conjugate subgradient theorem, 104
- convex feasibility problem, 208
- convex function, 21
- convex set, 3
- coordinate descent, 323
- coordinate-wise minimum, 407
- cyclic block proximal gradient method, 338
- cyclic shuffle, 346

- DBPG, *see* dual block proximal gradient
- decomposition method, 331
- denoising, 364
- descent direction, 195
- descent lemma, 109
- differentiable function, 48
- dimension, 2
- directional derivative, 44
- distance function, 22
- distributed optimization, 245
- domain, 14

- dot product, 4, 6
- DPG, 355
- dual block proximal gradient, 370
 - method, 369
- dual norm, 9
- dual projected subgradient method, 232
- dual proximal gradient, 355
- dual space, 9

- effective domain, 14
- eigenvalues, 182
- epigraph, 14
- ergodic convergence, 215
- Euclidean norm, 3
- Euclidean space, 3
- even function, 179
- exact line search, 196
- extended Moreau decomposition, 160
- extended real-valued functions, 13

- Farkas lemma, 28
- fast dual proximal gradient, 358
- fast proximal gradient method, 290
- FDPG, *see* fast dual proximal gradient
- Fejér monotonicity, 205
- Fenchel's dual, 102
- Fenchel's duality theorem, 102
- Fenchel's inequality, 88
- Fermat's optimality condition, 73
- Fermat–Weber problem, 75
- finite-dimensional vector space, 2
- first projection theorem, 147
- first prox theorem, 130
- FISTA, 290
- Fréchet differentiability, 48
- Frank Wolfe method, 379
- Fritz-John conditions, 81
- Frobenius norm, 6, 189
- functional decomposition method, 331

- generalized Cauchy–Schwarz, 9
- generalized conditional gradient method, 380
- generalized diagonal matrix, 188
- generalized permutation matrix, 180
- geometric median, 75
- global Lipschitz constant, 333
- gradient, 48
- gradient mapping, 272
- gradient method, 195
- greedy projection algorithm, 210

- Hadamard product, 5
- half-space, 3
- hard thresholding, 137
- hinge loss, 94
- Huber function, 163, 167, 169, 309
- hyperplane, 3

- identity transformation, 8
- incremental projected subgradient, 229
- indicator function, 14
- induced matrix norm, 7
- infimal convolution, 24, 102
- inner product, 2
- inner product space, 3
- ISTA, 271
- iterative shrinkage-thresholding algorithm, 271

- Jensen's inequality, 21

- KKT conditions, 81
- Kullback–Leibler divergence, 252
- Ky Fan norms, 190

- l_0 -norm, 19
- l_p -norm, 5
- l_∞ -norm, 5
- l_1 -regularized least squares, 295, 434
- L -smooth function, 107
- Lagrangian dual, 38
- level set, 15, 149
- line segment principle, 119
- linear approximation theorem, 112
- linear functional, 9
- linear programming, 212
- linear rate, 288
- linear transformation, 8
- linearly independent, 2
- log-sum-exp function, 98
- Lorentz cone, 154
- lower semicontinuous function, 15

- max formula, 47
- median, 73
- Minkowski sum, 26
- mirror descent, 247
- mirror-C method, 262
- Moore–Penrose pseudoinverse, 96
- Moreau decomposition, 160
- Moreau envelope, 163

- negative entropy, 96, 124
- negative sum of logs, 136, 152
- network utility maximization, 243

- non-Euclidean gradient method, 317
- non-Euclidean proximal gradient, 327
- non-Euclidean second prox theorem, 253
- nonnegative orthant, 5
- nonnegative part, 5
- norm, 2
- norm-dependent function, 180
- normal cone, 36
- nuclear norm, 189

- open ball, 2
- open line segment, 3
- orthogonal matrix, 6
- orthogonal projection, 49, 146

- partial conditional gradient norm, 402
- partial gradient mapping, 333
- partial prox grad mapping, 333
- permutation matrix, 180
- permutation symmetric function, 180
- polar cone, 27
- Polyak's stepsize, 204
- positive orthant, 5
- power method, 386
- primal counterpart, 316
- projected subgradient method, 201
- projected subgradient method, 202
- proper function, 14
- prox-grad operator, 271
- proximable, 432
- proximal gradient method, 269, 271
- proximal mapping, 129
- proximal point method, 288
- proximal subgradient method, 262

- \mathbf{Q} -inner product, 4
- \mathbf{Q} -norm, 4

- randomized block conditional gradient method, 402
- randomized block proximal gradient method, 348
- RBCG, *see* randomized block conditional gradient method
- RBPG, *see* randomized block proximal gradient method
- real vector space, 1
- relative interior, 43
- restarted FISTA, 299
- restarting, 299
- robust regression, 436

- S-FISTA, 310
- s -sparse vector, 174
- scalar, 1
- scalar multiplication, 1
- Schatten norm, 189
- second projection theorem, 157
- second prox theorem, 157
- singleton, 51
- singular value decomposition, 188
- Slater's condition, 82
- smoothness parameter, 107
- soft thresholding, 136, 142
- span, 2
- spectrahedron, 183
- spectral conjugate formula, 184, 190
- spectral decomposition, 182
- spectral function, 182, 189
- spectral norm, 7, 189
- standard basis, 4
- stationarity, 80
- steepest descent, 195
- stochastic projected subgradient method, 221
- strict separation theorem, 31
- strong convexity, 117
- strong duality theorem, 439
- strong subdifferential result, 39
- strongly convex set, 396
- subdifferentiable, 39
- subdifferential, 35
- subgradient, 35
- subgradient inequality, 35
- sublinear rate, 284
- sufficient decrease lemma, 272
- support function, 26, 161
- supporting hyperplane theorem, 41
- symmetric conjugate theorem, 181
- symmetric function, 179
- symmetric spectral function, 183, 189
- symmetric spectral set, 187, 194

- three-points lemma, 252
- total variation, 364
- trace norm, 189
- triangle inequality, 2
- two-sided soft thresholding, 151

- unbiased estimator, 222
- unit simplex, 5

- value function, 67
- variables decomposition method, 332
- vector space, 1
- von Neumann's trace inequality, 190

- weak subdifferential result, 39
- Wolfe's example, 197